AD-A259 784

JMENTATION PAGE

Form Approved
OMB No. 0704-0188

(2)

on is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, eting and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this lucing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1991	memorandum

4. TITLE AND SUBTITLE

Resolving Ambiguity in Nonmonotonic Inheritance Hierarchies

5. FUNDING NUMBERS

N00014-89-J-3202
N00014-85-K-0124

6. AUTHOR(S)

Lynn Andrea Stein

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Artificial Intelligence Laboratory
545 Technology Square
Cambridge, Massachusetts 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

AIM 1316

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
Information Systems
Arlington, Virginia 22217

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

DTIC
ELECTE
JAN 25 1993
S B

11. SUPPLEMENTARY NOTES

None

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution of this document is unlimited

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This paper describes a theory of inheritance theories. We present an original theory of inheritance in nonmonotonic hierarchies. The structures on which this theory is based delineate a framework that subsumes most inheritance theories in the literature, providing a new foundation for inheritance.

- Our path-based theory is sound and complete w.r.t. a direct model-theoretic semantics.
- Both the credulous and the skeptical conclusions of this theory are polynomial-time computable.
- We prove that true skeptical inheritance is not contained in the language of path-based inheritance.

(continued on back)

14. SUBJECT TERMS (key words)

inheritance
nonmonotonic logic
tractable inference
default reasoning

15. NUMBER OF PAGES

42

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED

Block 13 Continued:

Because our techniques are modular w.r.t. the definition of specificity, they generalize to provide a unified framework for a broad class of inheritance theories. By describing multiple inheritance theories in the same "language" of credulous extensions, we make principled comparisons rather than the *ad-hoc* examination of specific examples makes up most of the comparative inheritance work.

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

DTIC QUALITY INSPECTED 3

[12]

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 1316 August 1991

Resolving Ambiguity in Nonmonotonic Inheritance Hierarchies

Lynn Andrea Stein

Abstract

This paper describes a theory of inheritance theories. We present an original theory of inheritance in nonmonotonic hierarchies. The structures on which this theory is based delineate a framework that subsumes most inheritance theories in the literature, providing a new foundation for inheritance.

- Our path-based theory is sound and complete w.r.t. a direct model-theoretic semantics.
- Both the credulous and the skeptical conclusions of this theory are polynomial-time computable.
- We prove that true skeptical inheritance is not contained in the language of path-based inheritance.

Because our techniques are modular w.r.t. the definition of specificity, they generalize to provide a unified framework for a broad class of inheritance theories. By describing multiple inheritance theories in the same "language" of credulous extensions, we make principled comparisons rather than the *ad-hoc* examination of specific examples makes up most of the comparative inheritance work.

Copyright © Massachusetts Institute of Technology, 1991

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-85-K-0124.

93-01228

93 1 22 110

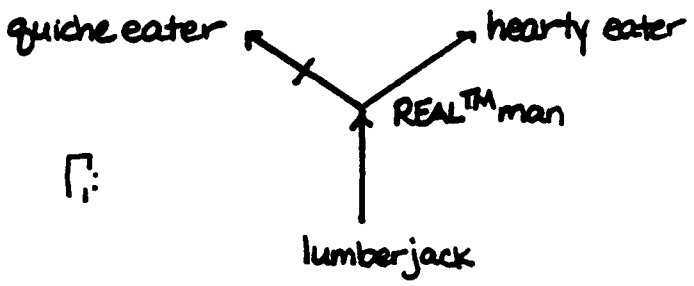


Figure 1: A simple inheritance hierarchy.

1 Introduction

Inheritance reasoning is ubiquitous. Hierarchies provide a concise encoding of much of our commonsense knowledge. They appear in various guises in the literature of artificial intelligence, but also throughout the various arts and sciences. Open nearly any textbook, and some hierarchy is likely to appear.

Reasoning with inheritance hierarchies is of particular interest to the artificial intelligence community because hierarchies are not only useful and native to human reasoning, but also simple and often tractable. The simplicity of the "inheritance problem" makes it attractive as a topic that can be considered in its entirety; its tractability makes it practical for the implementation of knowledge representation systems.

In this paper, we present several equivalent formulations of inheritance reasoning. Each approach may be useful in a particular context; by demonstrating their equivalence, we show that the most appropriate definition may be chosen. The techniques that we develop here generalize to a broad class of existing systems, providing a unified foundation for inheritance theory and a basis for comparative analysis, as well as extending previous results.

We begin, in section 2, by providing an intuitive description of what we intend by inheritance reasoning. We view inheritance hierarchies as representing primitive assertions in the knowledge base of some reasoning agent, and the inheritance problem as that of determining the agent's derived beliefs. The work in the remainder of this paper builds on this foundation.

In section 3, we give a formal path-based definition of inheritance. A path-based inheritance theory gives rules describing the admissible conclusions of a hierarchy. The theory presented here combines rules concerning the transitivity of primitive assertions with an ambiguity-resolution criterion to be invoked when two paths conflict. We describe the transitivity component by the notion of *reachability*; the resolution criterion is captured by *specificity*.

In section 4, we describe a model-theoretic semantics for inheritance hierarchies. A hierarchy defines a space of possible world-states—credulous extensions—or unambiguous interpretations. Specificity induces a preference relation over these world-states. In the resulting preferential semantics, each world-state has a classical model-theoretic interpretation, and the interpretation of an inheritance hierarchy is the set of models of preferred world-states. We demonstrate that the path-based theory of section 3 is sound and complete with respect to this model theory.

The model-theoretic semantics that we describe in section 4 differs from previous semantics for

inheritance hierarchies by separating the model preference criterion—specificity—from the semantic space. The language of credulous extensions is a common basis for the semantics of inheritance systems. By varying the definition of specificity—selecting different sets of preferred extensions—we obtain model-theoretic semantics for diverse inheritance theories. We present an alternate specificity criterion in section 8.

Section 5 describes a tractable algorithm for computing the inferences supported by an inheritance hierarchy. This is the first tractable theory for credulous inheritance—reasoning about what *might* be. Previous tractable inheritance theories have been limited to skeptical reasoning—computing what *must* be—or to an otherwise limited subset of the credulous conclusions.

In section 6, we present a reason maintenance labeling scheme for inheritance reasoners. We label each node of the hierarchy with a propositional formula corresponding to the conditions under which the associated inference—a is- or is-not-a x—holds. These labels keep track of all of the possible (or preferred) interpretations of an inheritance hierarchy simultaneously, allowing us to draw contingent conclusions and reason about the interrelatedness of inferences.

In section 7, we use the reason maintenance labeling to define and investigate the problem of ideally skeptical inheritance—computing the intersection of credulous extensions. Where credulous inheritance is analogous to propositional satisfiability, the ideally skeptical conclusions of a hierarchy are its theorems, or valid conclusions. We demonstrate that previous "skeptical" theories are not ideally skeptical, and prove that no purely path-based inheritance definition *can* compute the ideally skeptical conclusions of an inheritance hierarchy.

Section 8 and 9 examine several previous theories of inheritance. Section 8 replicates the work in the body of this paper for off-path inheritance. In section 9, we make more general comparisons with existing inheritance research and position our work in that larger context.

2 Hierarchies as Belief Spaces

In this section, we present an intuitive interpretation of inheritance hierarchies. In later sections, we give formal definitions of these ideas; here, we hope to motivate that more formal work by answering the question of "what we mean" when we draw an inheritance hierarchy. Regrettably, most previous theories of inheritance omit such a statement of intent, and the lack of such an intuitive semantics has been one of the criticisms levelled against the entire inheritance endeavor (*c.f.* Woods [42], Brachman [7], Bacchus [4], *etc.*).

Our interpretation of inheritance hierarchies is relatively simple. Each arc in an inheritance hierarchy—such as figure 1—represents an atomic assertion in the knowledge base of some reasoning agent—what this agent "believes," if you will.[1] Since this paper deals exclusively with defeasible inheritance, it is possible that an arc in the hierarchy—an atomic belief—is mistaken (e.g., *lumberjacks* might not, in fact, be *Real*™ men); however, in this agent's world-model, each of these atomic assertions holds.[2]

Reachability, or transitivity-by-default, poses a second constraint on world-models. This constraint arises when we try to apply an atomic assertion—*Real*™ men are *hearty* eaters—to some other class or individual—*lumberjacks*, or *Joe* the *Real*™ *man*. The fact that *Joe* may be a picky

[1] We put the word "believes" in quotation marks to emphasise that we are not proposing that edges of an inheritance hierarchy follow any realistic epistemic ontology; rather, we find the term belief, when removed from that formal context, to be suggestive of the kind of tentative assertion about the world that we wish to describe.

[2] The exception to this is the case in which the knowledge base contains both the atomic assertion that a (defeasibly) is an x, and the atomic assertion that a (defeasibly) is not an x. Here, there are (at least) two possibilities: the reasoner's beliefs may be ambiguous; or they may be (locally) inconsistent. We opt for the ambiguity interpretation (see below).

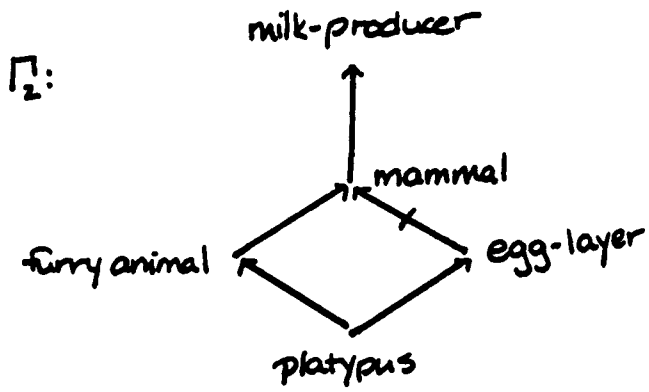


Figure 2: Is a *platypus* a *mammal*?

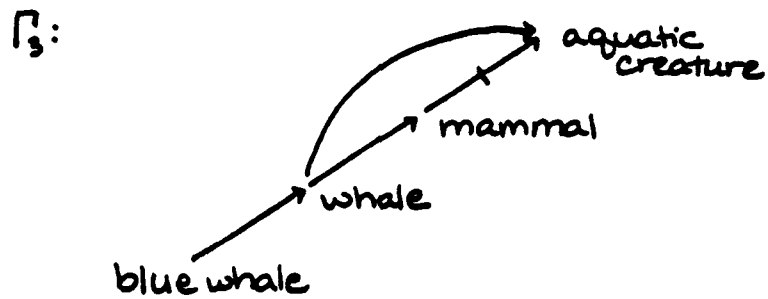


Figure 3: A *blue whale is* an *aquatic creature*.

eater does *not* defeat either the assertion that *Joe* is a *Real*™ *man*, or the assertion that *Real*™ *men* are generally *hearty eaters*; it merely makes *Joe* an atypical *Real*™ *man*. In the absence of conflicting information, the world-model of our reasoner is constrained so that subclasses and individuals are *typical* of their superclasses.

Finally, we have the third kind of defeasibility in inheritance hierarchies: the defeasibility of ambiguous conclusions. Unlike the defeasibility of atomic assertions or that of derived (but uncontested) conclusions, ambiguity's defeasibility arises when there are two explicit and conflicting arguments. In figure 2, the derived conclusion that *platypuses* are *mammals* is directly opposed by the (equally legitimate) conclusion that *platypuses* are *not mammals*. In this case, even the reasoning agent is assumed to be aware of the defeasibility. Indeed, we interpret such a hierarchy as asserting that the world must be in such a state that *either platypuses* are *mammals*, or they are not; but this reasoner does not know which. That is, the atomic assertions hold; and the assertions in some maximal consistent subset of their transitive closures hold; but there may be several such subsets, corresponding to several possible states of the world.

The situation here is not as hopeless as it may sound. First, the reasoner may prefer one of these possible world-states. For example, in figure 3, there are derived paths asserting that *blue whales*

are *aquatic creatures* (by virtue of their being *whales*, which are explicitly *aquatic*) and that *blue whales* are not *aquatic* (by virtue of their being *mammals*). If we were reasoning about *whales*, this conflict would be easily resolved: the assertion that *whales* are *aquatic* is explicit, and therefore blocks the derived assertion (through *mammals*) that *whales* are not *aquatic*. In the case of *blue whales*, we can resolve this ambiguity by resorting to arguments about specificity. Because *blue whales* are *mammals only* by virtue of being *whales*, information about *whales* is more specific to *blue whales* than information about *mammals*, and the reasoner prefers to believe that the actual world-state is one in which *blue whales* are *aquatic*. We discuss the notion of specificity, and the selection of preferred world-states, in greater detail below.

Second, even in truly ambiguous cases—such as Γ_2—the inheritance hierarchy may contain contingent information. In this case, the reasoner cannot determine whether the actual world corresponds to the possible world-state in which *platypuses* are *mammals*, or that in which they are not. However, the reasoner can assert that *if* the actual world-state is such that *platypuses are mammals*, then it follows that *platypuses* produce milk. In this fashion, the hierarchy supports conclusions relative to a particular possible world-state.

3 A Path-Based Definition

A path-based theory of inheritance is one which defines the admissible conclusions of an inheritance hierarchy to be precisely the set of conclusions supported by admissible paths in the hierarchy. For example, the reason that *lumberjacks* are believed to be *hearty eaters* (according to figure 1) is that there is a path from *lumberjack* through *Real*™ *man* to *hearty eater* in Γ_1. Like proof theories, path-based inheritance theories give almost algorithmic characterizations of the process by which conclusions are derived. In some cases, they lead to tractable inference procedures. In general, however, path-based inheritance theories lack model-theoretic semantics. Previous path-based theories of inheritance include the work of Touretzky, Horty, and Thomason [18, 17, 19, 38, 39, 40, 41], Sandewall [32], and Geffner and Verma [14].

In this section, we present a path-based theory of inheritance. In section 4 we describe a model-theoretic semantics for which this path-based theory is sound and complete. In section 5, we present an $O(n^4)$ algorithm for computing the conclusions of an inheritance hierarchy supported by this path-based definition.

Our path-based inheritance theory is similar to that of Touretzky [39, 40], save that it is *upwards reasoning*. Upwards inheritance reasons about the properties of some particular object, rather than about the set of objects possessing some particular property. Some ramifications of upwards *vs.* downwards inheritance are discussed by Touretzky *et al.* [41]; the computational advantages of upwards inheritance are described by Levesque and Selman [33] (see also section 5, below).

The approach that we describe in this section is also credulous: it allows a conclusion whenever that conclusion is consistent with *some* (preferred) interpretation of the hierarchy; or, whenever it is a plausible conclusion of the reasoner's explicit beliefs. Thus, in a hierarchy like Γ_2, we expect to draw both the conclusion that a *platypus* is-a *mammal*, and that a *platypus* is-not-a *mammal*. In section 6, we discuss the problem of *contingent reasoning*—determining the assumptions that underlie credulous conclusions. For example, we conclude that a *platypus* is-a *milk-producer* whenever we assume that it is-a *mammal*. In section 7, we discuss skeptical reasoning, in which only the certain—uncontested—conclusions of the hierarchy are considered.

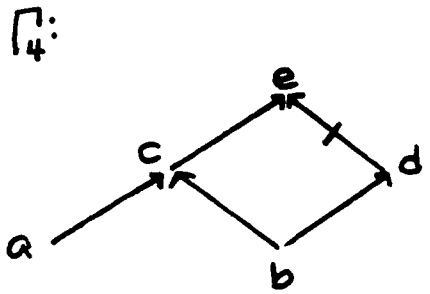


Figure 4: A hierarchy is ambiguous w.r.t. a focus node.

3.1 The Framework

An **inheritance hierarchy** $\Gamma = \langle V_\Gamma, E_\Gamma \rangle$ is a directed acyclic graph with positive and negative edges, intended to denote "is-a" and "is-not-a" respectively. We write a positive edge from a to x as $a \cdot x$, and a negative edge $a \cdot \neg x$. We call a sequence of positive edges $a \cdot s_1 \cdots s_n \cdot x$ (n $\geq$ 0),[3] a **positive path**, and a sequence of positive edges followed by a single negative edge $a \cdot s_1 \cdots s_n \cdot \neg x$ (n $\geq$ 0) a **negative path**.

A path, or **argument**, $a \cdot s_1 \cdots s_n \cdot (\neg)x$ supports the **inference** "a is (not) an x." We use the notation $a \rightarrow x$ (resp., $a \not\rightarrow x$) to stand for this inference, or conclusion, independently of the path through which it is derived. One inference—e.g., $a \rightarrow x$—may have many supporting arguments—$a \cdot s_1 \cdots s_n \cdot x$, $a \cdot t_1 \cdots t_m \cdot x$, *etc.*

Given an inheritance hierarchy $\Gamma = \langle V_\Gamma, E_\Gamma \rangle$ with nodes $a, x \in V_\Gamma$, we say that x is *reachable from* a (alternately, **a-reachable**) if there is some path $a \cdot s_1 \cdots s_n \cdot (\neg)x$ in E_Γ. If the final edge is positive—$s_n \cdot x$—we say that x is **positively reachable** from a; similarly, $s_n \cdot \neg x$ and **negatively a-reachable**. By extension, we say that an *edge* $s \cdot (\neg)x$ is reachable from a if s is positively a-reachable, and a *path* $s_1 \cdots s_n \cdot (\neg)x$ is reachable from a if every edge on that path is a-reachable. We say that a hierarchy Γ is **a-connected** if every node in V_Γ and every edge in E_Γ is reachable from a. When reasoning about an inheritance hierarchy w.r.t. a particular node, we call that node the **focus node**.

Ambiguity arises when two paths conflict. Formally, an inheritance hierarchy Γ is **ambiguous** w.r.t. a node a if there is some node $x \in V_\Gamma$ such that both $a \cdot s_1 \cdots s_n \cdot x$ and $a \cdot t_1 \cdots t_m \cdot \neg x$ are in E_Γ. In this case, we say that the ambiguity is *at* x. Ambiguity is always relative to a focus node: for example, Γ_4 is unambiguous w.r.t. a, but ambiguous w.r.t. b (at e).

Our definition of ambiguity differs from the conventional one. First, we introduce the notion of ambiguity w.r.t. a focus node. To previous theories, the hierarchy in figure 4 is simply ambiguous: w.r.t. a or b or c or d or e. Second, our definition of ambiguity is stronger than that in the literature. For example, by our definition, Γ_3 is ambiguous w.r.t. *whale*. This reflects the fact that there are paths *whale* · *aquatic creature* and *whale* · *mammal* · ¬*aquatic creature* in E_{Γ_3}. We shall see in the next section that specificity resolves this ambiguity and the *conclusions* derivable from Γ_3 are unambiguous. Previous inheritance theories do not distinguish between resolvable ambiguities—

[3] The notation $a \cdot s_1 \cdot s_2$ abbreviates the set of edges $\{a \cdot s_1, s_1 \cdot s_2\}$; $s_1 \cdots s_n$ abbreviates $\{s_1 \cdot s_2, s_2 \cdot s_3, \cdots s_{n-1} \cdot s_n\}$.

such as that in figure 3—and unambiguous hierarchies such as figure 1. Hierarchies with ambiguous conclusion sets—like figure 2—we term **truly ambiguous** w.r.t. focus nodes.

An inheritance hierarchy Γ **supports a path** $a \cdot s_1 \cdots s_n \cdot (\neg)x$, written $\Gamma \mid\!> a \cdot s_1 \cdots s_n \cdot (\neg)x$, if the corresponding sequence of edges $a \cdot s_1 \cdots s_n \cdot (\neg)x$ is in E_Γ *and it is admissible* according to specificity. We discuss admissibility in the next section. Γ **supports an inference** $a \rightarrow x$ (resp., $a \not\rightarrow x$) if it supports some corresponding path. For simplicity, we also allow the degenerate path a, with the corresponding inference $a \rightarrow a$.

3.2 Specificity

A **specificity criterion**, or preemption strategy, makes admissibility choices among certain competing paths. The idea of specificity dates from Touretzky's inferential distance [39, 40]. Since then, many definitions of specificity have appeared in the literature, but all operate on the same underlying principle: more specific information is more directly relevant. For example, in figure 3, information about *whales* is more specific to *blue whales* than information about *mammals*, so we can infer that *blue whales are aquatic*.

The framework that we have presented so far is compatible with many of the existing definitions of specificity. In this section, we present an original definition. Our specificity criterion closely resembles Touretzky's original notion of specificity [39, 40], but which is computable in polynomial time (see section 5). In section 8, we describe an alternate definition of specificity—one resembling that of Sandewall [32] and Horty *et al.* [18, 19]—and show how to integrate that definition into this framework.

An edge $v \cdot (\neg)x$ is **admissible** in Γ w.r.t. a if there is some path $a \cdot s_1 \cdots s_n \cdot v$, $(n \geq 0)$, in E_Γ, and

1. None of the edges of $a \cdot s_1 \cdots s_n \cdot v$ is redundant in Γ w.r.t. a,
2. Each of the edges of $a \cdot s_1 \cdots s_n \cdot v$ is admissible in Γ w.r.t. a, and
3. No intermediate node $a, s_1, \ldots s_n$ is a preemptor of $v \cdot (\neg)x$ w.r.t. a.

Intuitively, an edge is admissible if there is a non-redundant admissible path leading to it that contains no preempting intermediaries.

A **path is admissible** in Γ w.r.t. a if every edge in that path is admissible.

A node s is a **preemptor**[4] of $v \cdot x$ (resp., $v \cdot \neg x$) w.r.t. a if $s \cdot \neg x \in E_\Gamma$ (resp., $s \cdot x \in E_\Gamma$). For example, the positive edge from *whale* to *aquatic* preempts the negative edge from *mammal* to *aquatic*, w.r.t. both *whale* and *blue whale*.

The difficulties caused by redundant links were noted by Touretzky [39, 40]: Consider Γ_2 augmented by an additional edge from *blue whale* to *mammal*. This edge would be redundant: *blue whales* are typically *mammals* even without the explicit assertion. However, if the edge from *blue whale* to *mammal* were not excluded, there would be an admissible path *blue whale* · *mammal* · ¬*aquatic creature*—no intermediate node is a preemptor of *mammal* · ¬*aquatic creature*. Clearly, this is not the intended meaning here (or, indeed, in any network of this form, since the "*whale*" node is always more specific than the "*mammal*" node (w.r.t. "*blue whales*")).

An edge $b \cdot w$ is **redundant** in Γ w.r.t. focus node a if there is some positive path $b \cdot t_1 \cdots t_m \cdot w \in E_\Gamma$, $m \geq 1$, for which

[4] Perhaps "potential preemptor" would be a better term: if s is not on any (admissible, non-redundant, positive) path from a to v, then s effectively has no bearing on the admissibility of $v \cdot (\neg)x$. This is because preemptors are checked only in condition 3 of the definition of admissibility. In [35, 37, 36], we gave a stronger—and incorrect—condition here.

1. Each of the edges of $b \cdot t_1 \cdots t_m$ is admissible in Γ w.r.t. a,
2. There are no c and i such that $c \cdot \neg t_i$ is admissible in Γ w.r.t. a; there is no c such that $c \cdot \neg w$ is admissible in Γ w.r.t. a.[5]

Redundant edges may themselves be admissible. For example, an edge from *blue whale* to *mammal* in figure 3 would be admissible w.r.t. *blue whale*. However, redundant edges may not contribute to the admissibility of other edges: *mammal* · ¬*aquatic creature* is *not* admissible w.r.t. *blue whale*, in spite of the admissible path *blue whale* · *mammal*. Conditions 2 and 3 in the definition of admissibility hold for the path *blue whale* · *mammal* · *aquatic creature*, but condition 1 is violated.

The path-based definition of supports yields the following conclusions on the hierarchy Γ_2 in figure 2:

$\Gamma_2 \rhd$ *platypus*→*platypus*
$\Gamma_2 \rhd$ *platypus*→*furry animal*
$\Gamma_2 \rhd$ *platypus*→*egg-layer*
$\Gamma_2 \rhd$ *platypus*→*mammal*(***)
$\Gamma_2 \rhd$ *platypus*↛*mammal*(***)
$\Gamma_2 \rhd$ *platypus*→*milk-producer*

$\Gamma_2 \rhd$ *egg-layer*→*egg-layer*
$\Gamma_2 \rhd$ *egg-layer*↛*mammal*

$\Gamma_2 \rhd$ *furry animal*→*furry animal*
$\Gamma_2 \rhd$ *furry animal*→*mammal*
$\Gamma_2 \rhd$ *furry animal*→*milk-producer*

$\Gamma_2 \rhd$ *mammal*→*mammal*
$\Gamma_2 \rhd$ *mammal*→*milk-producer*

$\Gamma_2 \rhd$ *milk-producer*→*milk-producer*

In this case, specificity cannot resolve the ambiguity w.r.t. *platypus* at *mammal*, and Γ_2 supports both of the assertions marked (***): that *platypuses* are *mammals*, and that they are not *mammals*. In such a case—when a hierarchy *supports* conflicting paths—we say that the hierarchy is *truly ambiguous* w.r.t. *platypuses*.

The conclusion-set of Γ_3 is

$\Gamma_3 \rhd$ *blue whale*→*blue whale*
$\Gamma_3 \rhd$ *blue whale*→*whale*
$\Gamma_3 \rhd$ *blue whale*→*mammal*
$\Gamma_3 \rhd$ *blue whale*→*aquatic creature*

$\Gamma_3 \rhd$ *aquatic creature*→*aquatic creature*

$\Gamma_3 \rhd$ *whale*→*whale*
$\Gamma_3 \rhd$ *whale*→*mammal*
$\Gamma_3 \rhd$ *whale*→*aquatic creature*

$\Gamma_3 \rhd$ *mammal*→*mammal*
$\Gamma_3 \rhd$ *mammal*↛*aquatic creature*

In this case, specificity resolves the ambiguities w.r.t. *blue whale* and *whale* at *aquatic creature*. As a result,

$$\Gamma_3 \not\rhd \textit{blue whale} \not\to \textit{aquatic creature} \quad \text{and} \quad \Gamma_3 \not\rhd \textit{whale} \not\to \textit{aquatic creature}$$

The same set of conclusions would result if we added a redundant edge from *blue whale* to *mammal*.

4 Model-theoretic Semantics

The path-based inheritance theory of the previous section is a sort of "proof theory" for inheritance: given a hierarchy, it describes the rules by which conclusions may be drawn. In this section, we present a model-theoretic approach to understanding the meaning of an inheritance hierarchy. Rather than the admissible conclusions, this section speaks of the possible world-states—credulous extensions—that are "models" satisfying the constraints imposed by the hierarchy. In essence,

[5] Although the definitions of admissibility and redundancy are mutually dependent, they are not circular. Because the hierarchy is acyclic, it can be ordered topologically, and the definition of an admissible edge $v \cdot (\neg)x$ depends only on the redundancy of edges with endpoints strictly topologically earlier than x.

this section makes rigorous the informal characterization of the meaning of inheritance hierarchies described in section 2.

Previous theories of inheritance semantics have been *translational*: a hierarchy is expressed as a set of statements in some particular (nonmonotonic) logic, and the semantics of the logic provide semantics for the hierarchy. For example, McCarthy [27], Haugh [16], and Krishnaprasad, Kifer, and Warren [22, 23] translate inheritance hierarchies into particular circumscriptive theories [26]; Etherington and Reiter [11, 12, 13] treat hierarchies as theories of Reiter's default logic [31]; Przymusinska and Gelfond [15, 30] use Moore's autoepistemic logic [29] as a target language; Bacchus [4] bases his translation on a probabilistic logic; and Boutilier [6] uses a conditional logic.

While the semantics provided by most of these theories are as satisfying as the semantics of the underlying logics, they are less semantics of inheritance than theories of how inheritance relates to (or can be expressed in) those logics. In contrast, the theory that we present here is a theory of *direct semantics for inheritance hierarchies*. Hierarchies define a space of possible interpretations, or models. Specificity—drawn from the topological properties of the hierarchy—is used as a preference criterion over these models. The meaning of the hierarchy is the set of its maximally preferred models. The result is a preferential semantics like those of Bossu and Siegel [5], Etherington [10, 12], and Shoham [34] for more general nonmonotonic logics.

Further, virtually every previous—translational or path-based—approach to inheritance semantics contains a fixed ambiguity-resolving preemption strategy.[6] Although the selection of an appropriate preemption strategy is still a subject for debate in the inheritance literature (see, e.g., Touretzky *et al.*'s discussion of the "Clash of Intuitions" [41]), existing systems of inheritance assume some single, fixed strategy. Preemption strategies are variously embedded in the mechanics of path construction, the translation procedure, or the underlying nonmonotonic logic. This makes it extraordinarily difficult to compare underlying preemption strategies. As a result, most so-called "comparisons of inheritance theories" are in reality *ad hoc* comparisons of system performances on selected examples.

In contrast, our semantics is modular. We reduce a hierarchy to its possible interpretations, or credulous extensions. Our semantics for individual extensions is independent of the specificity criterion used to select among extensions. The credulous extension semantics generalize to arbitrary upwards inheritance theories, providing a common base for their semantics. By combining with different selection strategies, these extensions—possible world-states—give sound and complete model-theoretic semantics for alternate inheritance theories. This means we can compare the ambiguity-resolving heuristics of various inheritance theories directly and theoretically, rather than by resorting to *ad hoc* analysis of specific examples. We make use of this generality to explore off-path inheritance in section 8, below.

4.1 Semantics for Credulous Extensions

We focus first on the problem of providing a model-theoretic semantics for a single, unambiguous "credulous extension." In general, we believe that a translational approach to inheritance semantics is undesirable. Such an approach trades the topological information inherent in a hierarchy for the semantics for an existing logic. In particular, translational approaches provide less-than-satisfactory means for dealing with ambiguities. Often, they merely adopt the ambiguity-resolving strategy of the target logic, which may not be appropriate for inheritance hierarchies. Where they do provide additional, explicit ambiguity resolution, it is fixed as a part of the translation procedure.

The approach to semantics *for credulous extensions* that we present here is translational. However, it is not subject to these criticisms of translational approaches precisely because it provides

[6] The exception to this is Haugh [16]; see the discussion of his work in section 9, below.

translations only for credulous extensions—*unambiguous* subhierarchies—and not for a hierarchy as a whole. Thus, it does not rely on the translation procedure or the underlying logic for ambiguity-resolution strategies. All ambiguity-resolution is done in the (non-translational) process of *selecting* some preferred subset of the credulous extensions; once the set of credulous extensions—or possible interpretations—for a hierarchy has been established, the semantics for any single extension are straightforward. In the next section, we discuss the problem of deriving the appropriate preferences over credulous extensions.

A credulous extension corresponds to a possible world-state—one in the space of world-states defined by inheritance ambiguity. Formally, a **credulous extension** of an inheritance hierarchy Γ with respect to a node a is a maximal unambiguous a-connected subhierarchy of Γ with respect to a: if X_a^Γ is a credulous extension of Γ w.r.t. a, then for every edge $v \cdot (\neg)x \in E_\Gamma - E_{X_a^\Gamma}$, adding $v \cdot (\neg)x$ to X_a^Γ would make X_a^Γ ambiguous or not a-connected. An example of several credulous extensions—and some non-extensions—for the hierarchy of figure 2 is given in figure 5.

Several previous inheritance theories include a related but distinct concept. Touretzky [40] based his inheritance theory on a construct that he called a *grounded expansion*. Sandewall [32] used the term extension to refer to structures much like Touretzky's grounded expansions. However, our credulous extensions differ from Touretzky's grounded expansions and Sandewall's extensions—we will call them both *expansions*—in two significant ways. First, their expansions are *supersets* of the hierarchy—the original hierarchy *plus* some additional, "disambiguating" information. For example, in Γ_2, one expansion adds the additional edge *platypus* $\cdot$ *mammal*, while another adds *platypus* $\cdot \neg$*mammal*. In our terms, their expansions are themselves ambiguous hierarchies. Second, their expansions are of the hierarchy as a whole, while ours are w.r.t. a particular focus node. To Touretzky and Sandewall, Γ_2 simply has two expansions, period. In our theory, Γ_2 has two credulous extensions w.r.t. *platypus*, but only one w.r.t. *egg-layer* or *mammal*. In both of these ways, our credulous extensions correspond more closely to the extension of a predicate or of a default logic theory than to past definitions of the expansion of an inheritance hierarchy.

Because a credulous extension is unambiguous, every edge is admissible. Thus, instead of $X_a^\Gamma \mid\!> a \rightarrow x$—there is an admissible positive path from a to x—we need merely check that there is *some* positive path from a to x (resp., $a \not\rightarrow x$ and negative path). We make use of this to provide a straightforward model-theoretic semantics for a single extension:

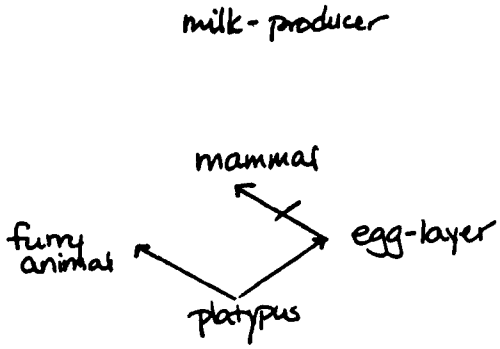
For every vertex $x \in V_{X_a^\Gamma}$, we create a unique propositional variable $\hat{x}$. $\widehat{X_a^\Gamma}$ is the theory (in the propositional calculus) given by

$$\hat{a} \qquad \bigwedge_{x \cdot y \in E_{X_a^\Gamma}} (\hat{x} \supset \hat{y}) \qquad \bigwedge_{x \cdot \neg y \in E_{X_a^\Gamma}} (\hat{x} \supset \sim \hat{y})^7$$

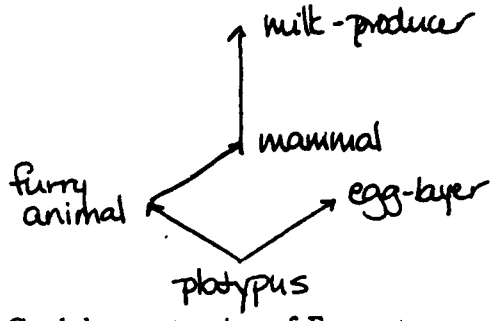
Since X_a^Γ is unambiguous, $\widehat{X_a^\Gamma}$ is consistent and has a model.

In other words, we translate the edges of X_a^Γ into material implications, allowing inference chains exactly when there are paths in the credulous extension. It would be nice if all of inheritance semantics were this easy. However, Thomason *et al.* demonstrate that the translation into propositional logic does not work even in the simpler case of strict (non-defeasible) inheritance [38], where local inconsistency in a hierarchy would lead to a globally inconsistent theory. In the general defeasible case, the propositional theory corresponding to an ambiguous inheritance hierarchy would be inconsistent. The only reason that the translation into a propositional theory works here is that it is the translation of a credulous extension: an *unambiguous* subhierarchy. Ambiguity resolution must therefore be applied in selecting the credulous extensions that are the preferred interpretations of the original hierarchy.

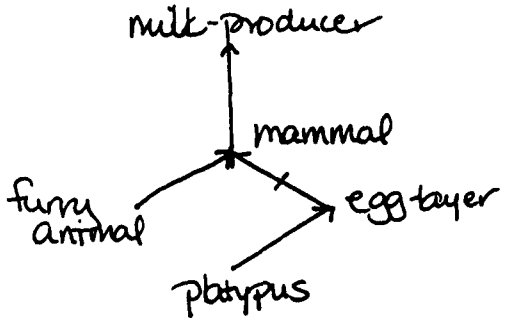
[7] $\wedge$, $\supset$, and $\sim$ should be read as propositional conjunction, material implication, and negation, respectively).



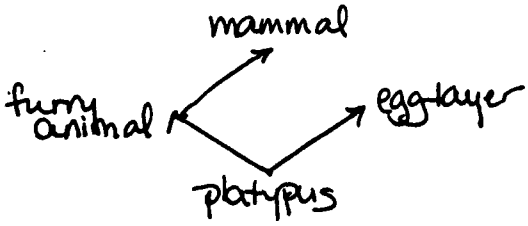
a. Credulous extension of Γ_2 w.r.t. *platypus*



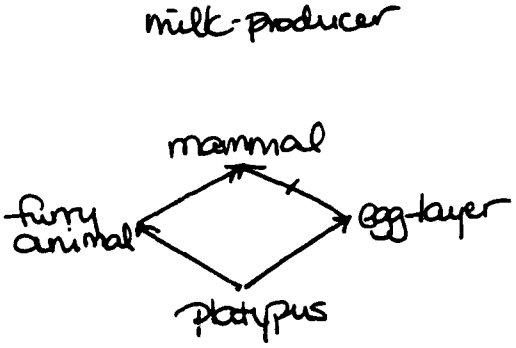
b. Credulous extension of Γ_2 w.r.t. *platypus*



c. Not *platypus*-connected



d. Can add *mammal* · *milk producer*



e. Ambiguous w.r.t. *platypus*

Figure 5: Some extensions and non-extensions of Γ_2.

Once we have chosen the preferred extensions—as we discuss in the next section—this translational semantics provides the desired result. The following theorem states that support—the existence of an admissible path—is equivalent to entailment in the propositional theory. Thus the path-based definition of support *for credulous extensions* is sound and complete w.r.t. the model-theoretic semantics defined here.

Theorem 1 (Soundness and Completeness for Credulous Extensions)
Let Γ be an inheritance hierarchy, with $a, x \in V_\Gamma$. Let X_a^Γ be a credulous extension of Γ w.r.t. a, and let $\widehat{X_a^\Gamma}$ be the propositional theory corresponding to X_a^Γ. The following are equivalent:

1. $X_a^\Gamma \mid\!\!> a \rightarrow x$ *(resp., $a \not\rightarrow x$).*
2. *x is positively a-reachable in X_a^Γ (resp., negatively a-reachable).*
3. $\widehat{X_a^\Gamma} \vdash \hat{x}$ *(resp., $\sim\hat{x}$).*
4. $\widehat{X_a^\Gamma} \models \hat{x}$ *(resp., $\sim\hat{x}$).*

This follows directly from the fact that every path in a credulous extension is admissible. Full proofs of all theorems may be found in appendix A.

4.2 Selecting Preferred Extensions

Given a single credulous extension, we have seen how to derive a model-theoretic interpretation of that extension. In fact, the semantics of the previous section would work any time we began with an unambiguous inheritance hierarchy (which would have only a single credulous extension w.r.t. any focus node). But inheritance hierarchies are generally ambiguous. As a result, Γ may have several extensions w.r.t. a given focus node, a. Some of these extensions may be more intuitive than others. In this section, we describe a means of selecting the *preferred* (more intuitive) extensions of Γ w.r.t. a. The semantics of Γ are then simply the sets of interpretations for the preferred extensions of Γ w.r.t. the nodes of V_Γ.

In inheritance hierarchies, *specificity* gives us a means of ruling out unintuitive interpretations. Thus, we use the definition of *admissibility* according to specificity, from section 3.2, above, to define the preference relation over credulous extensions. We say that one extension is *preferred* to another if it is "more consistent" with the constraints of specificity:

Let X_a^Γ and Y_a^Γ be two credulous extensions of an inheritance hierarchy Γ w.r.t. focus node a. Then **specificity prefers** X_a^Γ **to** Y_a^Γ ($X_a^\Gamma \preceq Y_a^\Gamma$) if there are some nodes v and x such that

1. X_a^Γ and Y_a^Γ agree on all edges whose endpoints topologically precede x,[8]
2. The edge $v \cdot (\neg)x$ is inadmissible in Γ w.r.t. a, and
3. Y_a^Γ contains that inadmissible edge: $\exists s_1, \ldots s_n$, $Y_a^\Gamma \mid\!\!> a \cdot s_1 \cdots s_n \cdot v \cdot (\neg)x$, and
4. X_a^Γ does not include it: $X_a^\Gamma \not\mid\!\!> a \cdot s_1 \cdots s_n \cdot v \cdot (\neg)x$. (Note that $X_a^\Gamma \mid\!\!> a \cdot s_1 \cdots s_n$, by 1.)

If a credulous extension is minimal under this preorder—i.e. no other extension is preferred to it—we call it a **preferred extension** of the hierarchy:

$$\mathcal{P}ref(\Gamma, a) = \{X_a^\Gamma \mid \forall Y_a^\Gamma, Y_a^\Gamma \not\preceq X_a^\Gamma\}$$

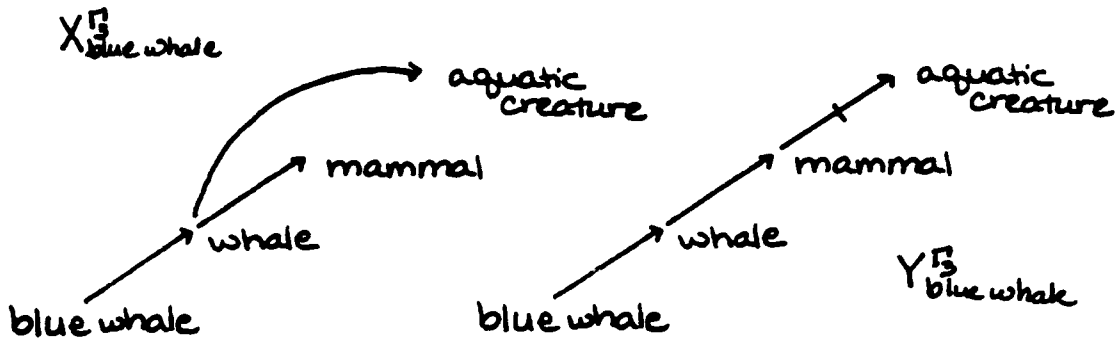


Figure 6: Specificity prefers $X^{\Gamma_3}_{\text{blue whale}}$ to $Y^{\Gamma_3}_{\text{blue whale}}$.

where X^{Γ}_{a} and Y^{Γ}_{a} are credulous extensions of Γ w.r.t. a.

Figure 6 shows the two credulous extensions of the hierarchy in figure 3. Specificity prefers $X^{\Gamma_3}_{\text{blue whale}}$ to $Y^{\Gamma_3}_{\text{blue whale}}$:

1. $X^{\Gamma_3}_{\text{blue whale}}$ and $Y^{\Gamma_3}_{\text{blue whale}}$ agree up to *aquatic creature*
2. *mammal* $\cdot$ $\neg$*aquatic creature* is inadmissible in Γ_3 w.r.t. *blue whale*
3. $Y^{\Gamma_3}_{\text{blue whale}} \rhd$ *blue whale* $\cdot$ *whale* $\cdot$ *mammal* $\cdot$ $\neg$*aquatic creature*, and
4. $X^{\Gamma_3}_{\text{blue whale}} \not\rhd$ *blue whale* $\cdot$ *whale* $\cdot$ *mammal* $\cdot$ $\neg$*aquatic creature*.

Theorem 2 is a soundness and completeness theorem for the path-based definition of section 3 w.r.t. the complete model-theoretic semantics defined here: selecting preferred extensions and providing translational semantics for those preferred extensions. In section 8, we demonstrate that other inheritance theories, with other definitions of specificity, can be described in these terms. This enables us to use the definitions of this section and section 3 to give similar soundness and completeness results for several existing inheritance theories.

Theorem 2 (Soundness and Completeness)
Let Γ be an inheritance hierarchy, with $a, x \in V_\Gamma$, and let $\mathcal{P}ref(\Gamma, a)$ be the set of preferred extensions of Γ w.r.t. a. Then $\Gamma \rhd a \rightarrow x$ iff there is some preferred extension $X^{\Gamma}_{a} \in \mathcal{P}ref(\Gamma, a)$ such that $\widehat{X^{\Gamma}_{a}} \models \widehat{x}$ (resp., $a \not\rightarrow x$ and $\sim\widehat{x}$).[9]

The complete model-theoretic interpretation of Γ_2 in figure 2, is therefore the set of models of preferred extensions. In this case, specificity cannot disambiguate the hierarchy, and both extensions are preferred. The propositional theories w.r.t. *platypus* are

$$\widehat{platypus} \;\wedge\; (\widehat{platypus} \supset \widehat{furry\ animal}) \;\wedge\; (\widehat{furry\ animal} \supset \widehat{mammal}) \;\wedge\; (\widehat{mammal} \supset \widehat{milk\text{-}producer}) \;\wedge\; (\widehat{platypus} \supset \widehat{egg\text{-}layer})$$

[8] That is, if every topological sort places s and t before x, then X^{Γ}_{a} and Y^{Γ}_{a} agree on edges $s \cdot (\neg)t$.

[9] Since $\Gamma \rhd a \rightarrow x$ whenever *some* preferred extension (of Γ w.r.t. a) entails $\widehat{x}$, $\rhd$ is analogous to propositional satisfiability. This is the essence of credulous inheritance: a conclusion is admissible if there is some (preferred) possible world-state supporting it. In section 7, we discuss the problem of skeptical inheritance—"valid" conclusions—entailed by all preferred possible world-states. In terms of the framework of this section, these are the conclusions supported by *all* preferred credulous extensions.

entailing the conclusions

$$\widehat{platypus} \qquad \widehat{furry\ animal} \qquad \widehat{mammal} \qquad \widehat{milk\text{-}producer} \qquad \widehat{egg\text{-}layer}$$

and

$$\widehat{platypus} \ \wedge \ (\widehat{platypus} \supset \widehat{furry\ animal}) \ \wedge \ (\widehat{platypus} \supset \widehat{egg\text{-}layer}) \\ \wedge \ (\widehat{egg\text{-}layer} \supset \sim\widehat{mammal})$$

which entails

$$\widehat{platypus} \qquad \widehat{furry\ animal} \qquad \sim\widehat{mammal} \qquad \widehat{egg\text{-}layer}$$

The preferential semantics allows the conclusion that *platypus* is-a x if x is entailed by the models of either of these theories, so *playtpus* is-a

$$platypus \qquad furry\ animal \qquad mammal \qquad milk\text{-}producer \qquad egg\text{-}layer$$

and *platypus* is-not-a *mammal*. The hierarchy is unambiguous w.r.t. its other nodes. The conclusions w.r.t. *furry animal* are $\widehat{furry\ animal}$, $\widehat{mammal}$, and $\widehat{milk\text{-}producer}$, since $\widehat{X^{\Gamma_2}_{furry\ animal}}$ is

$$\widehat{furry\ animal} \ \wedge \ (\widehat{furry\ animal} \supset \widehat{mammal}) \ \wedge \ (\widehat{mammal} \supset \widehat{milk\text{-}producer})$$

$\widehat{X^{\Gamma_2}_{mammal}}$ is

$$\widehat{mammal} \ \wedge \ (\widehat{mammal} \supset \widehat{milk\text{-}producer})$$

which entails $\widehat{mammal}$ and $\widehat{milk\text{-}producer}$, and $\widehat{X^{\Gamma_2}_{milk\text{-}producer}}$ is $\widehat{milk\text{-}producer}$. $\widehat{X^{\Gamma_2}_{egg\text{-}layer}}$ is

$$\widehat{egg\text{-}layer} \ \wedge \ (\widehat{egg\text{-}layer} \supset \sim\widehat{mammal})$$

entailing $\widehat{furry\ animal}$ and $\sim\widehat{mammal}$.

In contrast, specificity does provide a preference over the credulous extensions of the hierarchy in figure 3 w.r.t. *blue whale* and *whale*. The interpretation of Γ_3 w.r.t. *blue whale* is therefore the interpretation of the single preferred extension, $\widehat{X^{\Gamma_3}_{blue\ whale}}$, from figure 6:

$$\widehat{blue\ whale} \ \wedge \ (\widehat{blue\ whale} \supset \widehat{whale}) \ \wedge \ (\widehat{whale} \supset \widehat{mammal}) \\ \wedge \ (\widehat{whale} \supset \widehat{aquatic\ creature})$$

This entails the conclusions (w.r.t. *blue whale*)

$$\widehat{blue\ whale} \qquad \widehat{whale} \qquad \widehat{mammal} \qquad \widehat{aquatic\ creature}$$

Similarly, specificity prefers the extension w.r.t. *whale* in which *whales* are *aquatic*, with the corresponding propositional theory

$$\widehat{whale} \ \wedge \ (\widehat{whale} \supset \widehat{mammal}) \ \wedge \ (\widehat{whale} \supset \widehat{aquatic\ creature})$$

and the hierarchy entails the inferences that a *whale* is-a *whale*, *mammal*, and *aquatic creature*. The hierarchy is unambiguous w.r.t. *mammals* and *aquatic creatures*, so there is only one credulous extension w.r.t. each of these:

$$\widehat{mammal} \ \wedge \ (\widehat{mammal} \supset \sim\widehat{aquatic\ creature})$$

and a *mammal* is-a *mammal*, but is-not-a *aquatic creature*; and

$$\widehat{aquatic\ creature}$$

so an *aquatic creature* is-a *aquatic creature*.

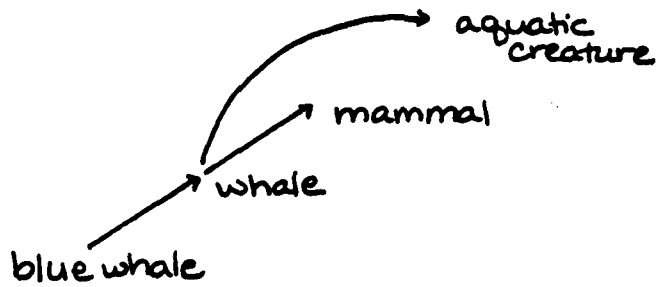


Figure 7: $\Sigma^{\Gamma_3}_{blue\ whale}$ computes specificity on Γ_3 w.r.t. *blue whale*.

5 Computing Specificity

The preference criterion approach to ambiguity resolution is satisfying from a semantic perspective, but it does not tell us much about how to *derive* the preferred extensions of a hierarchy. In this section, we give a polynomial-time—$O(n^5)$—algorithm for computing Σ^{Γ}_{a}, the *specificity extension* of Γ w.r.t. a. This algorithm demonstrates that upwards inheritance is tractable for credulous (goal-directed) on-path reasoning. In section 7, we extend this result to include ideally skeptical reasoning; in section 8, we give a similar algorithm for off-path inheritance.

There are several other computational theories of inheritance in the literature. Those of Horty *et al.* [18, 19] and Haugh [16] as well as the ambiguity-propagating algorithm of our [36, section 4] are "skeptical"—they attempt to compute only those conclusions that are unopposed. We discuss skeptical theories in section 7, below, where we show that none of these previous theories are ideally skeptical. Geffner and Verma [14] also offer an algorithm that is sound but incomplete for their inheritance theory. No previous path-based theory has been shown tractable for credulous inheritance.

Etherington [12, p. 89] gives an algorithm for finding extensions of general default theories. Kautz and Selman [21, theorem 1] have adapted this algorithm to find a single (arbitrary) extension of a disjunction-free ordered default theory. This class includes the inheritance theory of Etherington and Reiter [12, 13]. However, Kautz and Selman's $O(n^2)$ algorithm finds only a single arbitrary extension of the hierarchy, and cannot be used to determine whether a particular conclusion is supported. Indeed, Kautz and Selman demonstrate that goal-directed reasoning—determining whether there is an extension supporting some particular conclusion—is, for ordered default theories, $\mathcal{NP}$-hard [21, theorem 3]. Similarly, they show that skeptical reasoning for the same class of theories is $\mathcal{NP}$-hard [21, theorem 7]. In [33, theorem 2], Levesque and Selman demonstrate that downward inheritance reasoning (as espoused, e.g., by Touretzky [39, 40]) is $\mathcal{NP}$-hard, whether skeptical or credulous, on- or off-path.

Our algorithm is therefore the first sound and complete algorithm for credulous inheritance reasoning. It generates Σ^{Γ}_{a} the **specificity extension** of Γ w.r.t. a. Σ^{Γ}_{a} is a subhierarchy of Γ containing only and exactly the admissible edges of Γ w.r.t. a. For example, the specificity extension of the hierarchy in figure 3 w.r.t. *blue whale* is shown in figure 7.

The following algorithm always yields a unique specificity extension Σ^{Γ}_{a} for a hierarchy Γ w.r.t.

focus node a.

COMPUTE-SPECIFICITY-EXTENSION (Γ, a)

Let Σ_a^Γ contain the a-reachable nodes of Γ, and no edges.	; *Σ_a^Γ will be the specificity extension of Γ w.r.t. a.*
$R := \emptyset$	; *R will hold the redundant edges of Γ w.r.t. a.*
For each node x in Σ_a^Γ, in topological order	; *From a to ...:*
Add edges $v \cdot (\neg)\, x$ from Γ to Σ_a^Γ	; *Restore x-edges*
For each edge $v \cdot (\neg)\, x$ in Σ_a^Γ, in reverse topological order	; *Now weed out preempted x-edges:*
Let Σ^* be Σ_a^Γ minus any nodes that are preemptors of $v \cdot (\neg)\, x$ (i.e. remove any edges into and out of preempting nodes)	; *Remove potential preemptors*
If Σ^* no longer contains a positive path from a to v then remove the edge $v \cdot (\neg)\, x$ from Σ_a^Γ	; *Is v still positively a-reachable?* ; *If not, it's preempted*
For each remaining positive edge $p \cdot x$	; *Check if it's redundant*
If there is a positive path from p to x in Σ_a^Γ that goes through no negatively a-reachable nodes	; *If so,*
then move $p \cdot x$ from $E_{\Sigma_a^\Gamma}$ to R	; *Remove the edge—temporarily*
Restore the redundant edges in R to $E_{\Sigma_a^\Gamma}$	; *Redundant edges are still admissible*
Return Σ_a^Γ	

Theorem 3 (Complexity of COMPUTE-SPECIFICITY-EXTENSION)
Constructing Σ_a^Γ is $O(n^5)$

The following theorem is essentially a proof of correctness for COMPUTE-SPECIFICITY-EXTENSION. It says that Σ_a^Γ contains all and only those edges of Γ that are admissible w.r.t. a.

Theorem 4 (Correctness of COMPUTE-SPECIFICITY-EXTENSION)
Let Γ be an inheritance hierarchy, with $a \in V_\Gamma$, and let Σ_a^Γ be the specificity extension of Γ w.r.t. a. Then $E_{\Sigma_a^\Gamma} = \{v \cdot (\neg)x \mid v \cdot (\neg)x$ admissible in Γ w.r.t. $a\}$

The subhierarchy of admissible (w.r.t. a) edges in Γ may still be ambiguous. For example, all of the edges in a diamond ambiguity such as the platypus diamond of figure 2 are admissible; $\Sigma_{platypus}^{\Gamma_2}$, the specificity extension of Γ_2 w.r.t. *platypus*, is simply Γ_2 and, like Γ_2, has two credulous extensions w.r.t. *platypus*. However, the ambiguities that *can* be resolved by specificity have been: the credulous extensions of Σ_a^Γ are precisely the preferred credulous extensions of Γ w.r.t. a.

Corollary 4.1 *Let Γ be an inheritance hierarchy, with $a \in V_\Gamma$. Let $\mathcal{P}ref(\Gamma, a)$ be the set of preferred credulous extensions of Γ w.r.t. a and let Σ_a^Γ be the specificity extension of Γ w.r.t. a. Then $X_a^\Gamma \in \mathcal{P}ref(\Gamma, a)$ iff X_a^Γ is a credulous extension of Σ_a^Γ w.r.t. a.*

Corollary 4.2 *Let Γ be an inheritance hierarchy, with $a \in V_\Gamma$. Let Σ_a^Γ be the specificity extension of Γ w.r.t. a. Then $\Gamma \mid\!> a \rightarrow x$ iff x is positively a-reachable in Σ_a^Γ (resp., $a \not\rightarrow x$ and negatively a-reachable).*

This means that an inference $a \rightarrow x$ is supported by Γ—alternately, a preferred credulous extension of Γ w.r.t. a entails $\hat{x}$—iff x is positively a-reachable in Σ_a^Γ. Since verifying reachability is linear in the number of edges, it follows that computing the credulous conclusions of Γ is also tractable.

Corollary 4.3 (Complexity of Credulous Inheritance)
Deciding whether $\Gamma \mid\!> a \rightarrow x$ is $O(n^5)$ (resp., $a \not\rightarrow x$).

6 Reason Maintenance and Inheritance

In this section, we define a reason maintenance system for inheritance hierarchies. A reason maintenance system is a construction that keeps track of the assumptions or justifications supporting a particular conclusion; several systems of this type appear in the literature [8, 9, 24, 25, 28].

In the case of inheritance hierarchies, the reason maintenance system keeps track of the multiple credulous extensions of a hierarchy (w.r.t. each node), using a set of propositional labels for the nodes of that hierarchy. This function mimics the behavior of de Kleer's assumption-based truth maintenance system (*ATMS*) [8]; when we introduced the labeling scheme in [35], we called it the *ATMS-labeling* of the hierarchy.

The reason maintenance labeling keeps track of all of the possible interpretations—the credulous extensions—at once. We exploit this feature to draw contingent conclusions: that is, to determine what follows *if* we make a certain assumption. For example, in figure 2, we can conclude that *if platypuses* are *mammals*, *then* they *produce milk*. In section 7, below, we use the labeling to examine the problem of *skeptical* inheritance—computing what conclusions hold in every credulous extension.

For each pair of nodes, a and x, in V_Γ we define two labels: $[\![x]\!]_a^\Gamma$, the conditions under which a is an x, and $[\![\overline{x}]\!]_a^\Gamma$, the conditions under which a is not an x. $[\![\ ]\!]_a^\Gamma$ and $[\![\overline{\ }]\!]_a^\Gamma$ may be thought of as operators on nodes returning boolean formulae. For example, we would expect that $[\![\textit{hearty eater}]\!]_{\textit{lumberjack}}^{\Gamma_1}$ would be $[\top]$—*lumberjacks* are expected to be *hearty eaters*—while the *negative* label $[\![\overline{\textit{quiche-eater}}]\!]_{\textit{lumberjack}}^{\Gamma_1}$ would be $[\top]$—*lumberjacks* normally don't *eat quiche*. However, $[\![x]\!]_a^\Gamma$ isn't always $\sim[\![\overline{x}]\!]_a^\Gamma$: both the positive label $[\![\textit{quiche-eater}]\!]_{\textit{hearty eater}}^{\Gamma_1}$ and the negative label $[\![\overline{\textit{quicheeater}}]\!]_{\textit{hearty eater}}^{\Gamma_1}$ are $[\bot]$, since Γ_1 gives us no information about whether *hearty eaters* tend to *eat quiche*.

In general, we might reason about whether a is an x in some particular Γ as follows:

> Let $\mathcal{P}os_\Gamma(x)$ be the "positive children" of x in Γ—the nodes $p_i \in V_\Gamma$ with positive edges $p \cdot x \in E_\Gamma$—and let $\mathcal{N}eg_\Gamma(x)$ be the "negative children" of x in Γ—$n \in V_\Gamma \mid n \cdot \neg x \in E_\Gamma$. Suppose that, for all $p \in \mathcal{P}os_\Gamma(x)$, a is not a p; i.e., $\Gamma \not\mid\!> a \rightarrow p$. Then $\Gamma \not\mid\!> a \rightarrow x$, because there is no support for x. So
>
> $$\sim[\![\mathcal{P}os_\Gamma(x)]\!]_a^\Gamma \quad \supset \quad \sim[\![x]\!]_a^\Gamma$$
>
> Here, we have introduced a notational shorthand: $[\![\mathcal{P}os_\Gamma(x)]\!]_a^\Gamma$ really means $(\bigvee_{p \in \mathcal{P}os_\Gamma(x)} [\![p]\!]_a^\Gamma)$, where an empty disjunction is to be read as $\bot$.

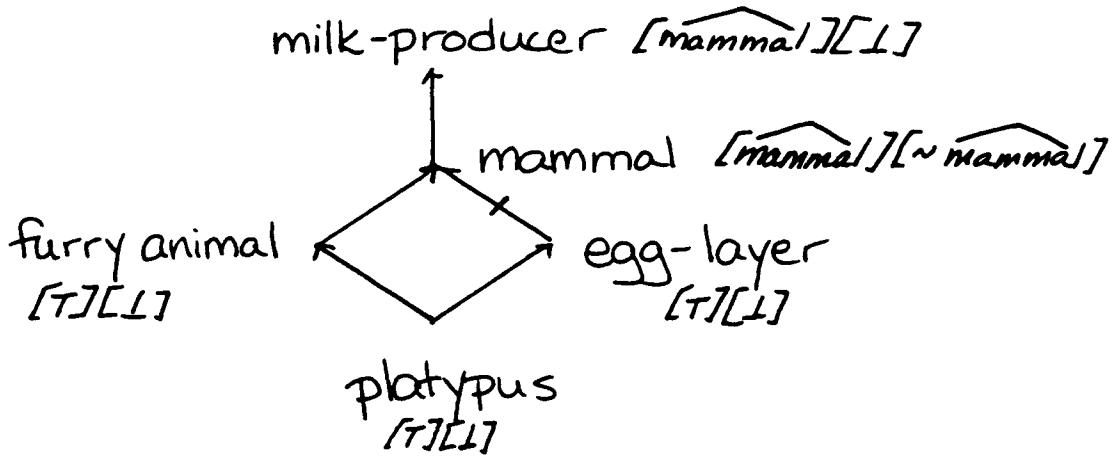


Figure 8: The labeling of Γ_2 w.r.t. *platypus*.

On the other hand, if a is a p, for some $p \in \mathcal{P}os_\Gamma(x)$, but a is definitely not an n for any of x's negative children n, then a is also an x: x's positive children provide it with a supporting argument, and there are no explicit counterarguments. That yields

$$[\![\mathcal{P}os_\Gamma(x)]\!]^\Gamma_a \wedge (\sim[\![\mathcal{N}eg_\Gamma(x)]\!]^\Gamma_a) \quad \supset \quad [\![x]\!]^\Gamma_a$$

Finally, it could be that at least one of $\mathcal{P}os_\Gamma(x)$ is true, but at least one of $\mathcal{N}eg_\Gamma(x)$ is also true. In this case, x is ambiguous: it has some support, but there is also a counterargument. If x is ambiguous, its value is unconstrained. In some credulous extensions, we would expect $[\![x]\!]^\Gamma_a$ to be true, and in others it should be false. When this situation arises, x becomes a choice-point and we introduce $\widehat{x}$—the free variable corresponding to the node x—into the labeling scheme. The truth value assigned to $\widehat{x}$ shouldn't matter most of the time, but in this special circumstance—when $[\![\mathcal{P}os_\Gamma(x)]\!]^\Gamma_a$ and $[\![\mathcal{N}eg_\Gamma(x)]\!]^\Gamma_a$, so x is ambiguous—we want $[\![x]\!]^\Gamma_a$ to vary freely with $\widehat{x}$.

$$([\![\mathcal{P}os_\Gamma(x)]\!]^\Gamma_a \wedge [\![\mathcal{N}eg_\Gamma(x)]\!]^\Gamma_a) \quad \supset \quad ([\![x]\!]^\Gamma_a \equiv \widehat{x})$$

Conjoining these three constraints gives us the conditions under which a is an x. The positive label for x w.r.t. a is given by:

$$[\![x]\!]^\Gamma_a \quad \equiv \quad [([\![\mathcal{P}os_\Gamma(x)]\!]^\Gamma_a \wedge (\sim[\![\mathcal{N}eg_\Gamma(x)]\!]^\Gamma_a)) \vee ([\![\mathcal{P}os_\Gamma(x)]\!]^\Gamma_a \wedge [\![\mathcal{N}eg_\Gamma(x)]\!]^\Gamma_a \wedge \widehat{x})]$$

whenever x is not a leaf.

In addition, we have $[\![a]\!]^\Gamma_a = [\top]$—$a$ is always an a—and $\forall x$, if $x \neq a$ is a leaf of the hierarchy, $[\![x]\!]^\Gamma_a = [\bot]$—trivially, $\Gamma \not\rhd a \rightarrow x$. These rules for generating labels mimic the construction of paths—reachability—by concatenating the edges of the hierarchy.

There is a similar labeling scheme for $[\![\overline{x}]\!]^\Gamma_a$, the conditions under which $\Gamma \rhd a \not\rightarrow x$:

1. $[\![\overline{a}]\!]^\Gamma_a = [\bot]$: a is never a non-a.

2. If $x \neq a$ is a leaf of Γ, $[\![\overline{x}]\!]^\Gamma_a = [\bot]$: since x cannot be a-reachable, $\Gamma \not\rhd a \not\rightarrow x$.

3. For a generic node, x,

$$[\![\overline{x}]\!]^\Gamma_a \quad \equiv \quad ((\sim[\![\mathcal{P}os_\Gamma(x)]\!]^\Gamma_a) \wedge [\![\mathcal{N}eg_\Gamma(x)]\!]^\Gamma_a) \vee ([\![\mathcal{P}os_\Gamma(x)]\!]^\Gamma_a \wedge [\![\mathcal{N}eg_\Gamma(x)]\!]^\Gamma_a \wedge \sim\widehat{x})$$

For example, the labeling of the *platypus* diamond of figure 2 w.r.t. *platypus* is shown in figure 8. In this hierarchy, $[\![mammal]\!]^{\Gamma_2}_{platypus} = [\widehat{mammal}]$. Although this label says very little—*platypuses* are *mammals* in those extensions in which they are *mammals*—the propagation of labels yields $[\![milk\text{-}producer]\!]^{\Gamma_2}_{platypus} = [\widehat{mammal}]$—*platypuses* are also *milk-producers* in precisely those extensions in which they are *mammals*. This is exactly the sort of contingent reasoning that we would expect our interpretation of inheritance hierarchies to capture.

Theorem 5 demonstrates that these labels keep track of all of the credulous extensions of a hierarchy at once. In the following sections, we present several useful corollaries of this theorem.

Theorem 5 (Correspondence Theorem)
Let Γ be an inheritance hierarchy, with $a, x \in V_\Gamma$. Then there is a correspondence between the set of credulous extensions of Γ w.r.t. a, and the truth-assignments to the free variables in the labels of Γ w.r.t. a, such that the truth-assignment assigns $[\![x]\!]^\Gamma_a = [\top]$ iff the corresponding extension supports $a \rightarrow x$ (resp., $[\![\overline{x}]\!]^\Gamma_a$ and $a \not\rightarrow x$).

The mapping from extensions of Γ w.r.t. a to truth-assignments assigns a variable $\widehat{x}$ $[\top]$ iff a is-a x in the extension; the inverse mapping creates an extension containing those edges $q \cdot s \in E_\Gamma$ for which $[\![q]\!]^\Gamma_a = [\![s]\!]^\Gamma_a = [\top]$, and $q \cdot \neg s \in E_\Gamma$ whenever $[\![q]\!]^\Gamma_a = [\top]$ and $[\![\overline{s}]\!]^\Gamma_a = [\bot]$. A proof of the Correspondence Theorem may be found in appendix A.

It follows immediately that the labels of an inheritance hierarchy are computing reachability:

Corollary 5.1 *Let Γ be an inheritance hierarchy, with $a, x \in V_\Gamma$. Then $[\![x]\!]^\Gamma_a$ is satisfiable iff x is positively a-reachable in Γ. (resp., $[\![\overline{x}]\!]^\Gamma_a$ and negatively a-reachable).*

Of course, reachability is not really what we are interested in. Γ supports an inference if its conclusion is reachable by an *admissible* path, not just any path. That is, the labels are not taking into account specificity—the preemption of an argument by a more specific counterargument.

Happily, we have a way of constructing the subhierarchy of admissible edges in Γ w.r.t. a—this is Σ^Γ_a. It follows immediately from theorem ?? and corollary 5.1 that satisfiability of the labels of Σ^Γ_a w.r.t. a corresponds to admissibility in Γ w.r.t. a, i.e. $\mid\!>$.

Corollary 5.2 *Let Γ be an inheritance hierarchy, with $a, x \in V_\Gamma$, and let Σ^Γ_a be the specificity extension of Γ w.r.t. a. Then $[\![x]\!]^{\Sigma^\Gamma_a}_a$ is satisfiable iff $\Gamma \mid\!> a \rightarrow x$ (resp., $[\![\overline{x}]\!]^{\Sigma^\Gamma_a}_a$ and $a \not\rightarrow x$).*

This means that to take specificity into account—to reason only about Γ's *preferred* extensions w.r.t. a—we can simply apply COMPUTE-SPECIFICITY-EXTENSION to Γ, a and label the resulting Σ^Γ_a. Any contingent conclusions—or, as we shall see in the next section, skeptical conclusions—that we draw using the labels of Σ^Γ_a w.r.t. a apply to Γ *taking specificity into account.*

An alternative but more complex approach to labeling with specificity might be to integrate the specificity criterion directly into the labeling scheme. In this case, we would condition the acceptance of a node on its *non-preempted* children. Unfortunately, even for a relatively simple preemption scheme such as the one described in section 3.2, these labels are quite complicated. For

example, the "specificity label" for $a \rightarrow x$ is

$$\left[\begin{array}{l} \left(\begin{array}{l} \bigvee_{p \in \mathcal{C}pos_\Gamma(x)} \left([\![p]\!]_a^\Gamma \wedge \sim \left(\bigwedge_{\{k \mid a \cdot s_1^k \cdots s_n^k \cdot p \text{ is admissible}\}} \bigvee_{\{i \mid s_i^k \cdot \neg x \in E_\Gamma\}} [\![s_i^k]\!]_a^\Gamma \right) \right) \\ \wedge \sim \left(\bigvee_{n \in \mathcal{C}neg_\Gamma(x)} [\![n]\!]_a^\Gamma \wedge \sim \left(\bigwedge_{\{k \mid a \cdot s_1^k \cdots s_n^k \cdot n \text{ is admissible}\}} \bigvee_{\{i \mid s_i^k \cdot x \in E_\Gamma\}} [\![s_i^k]\!]_a^\Gamma \right) \right) \end{array}\right) \\ \vee \left(\begin{array}{l} \bigvee_{p \in \mathcal{C}pos_\Gamma(x)} \left([\![p]\!]_a^\Gamma \wedge \sim \left(\bigwedge_{\{k \mid a \cdot s_1^k \cdots s_n^k \cdot p \text{ is admissible}\}} \bigvee_{\{i \mid s_i^k \cdot \neg x \in E_\Gamma\}} [\![s_i^k]\!]_a^\Gamma \right) \right) \\ \wedge \bigvee_{n \in \mathcal{C}neg_\Gamma(x)} \left([\![n]\!]_a^\Gamma \wedge \sim \left(\bigwedge_{\{k \mid a \cdot s_1^k \cdots s_n^k \cdot n \text{ is admissible}\}} \bigvee_{\{i \mid s_i^k \cdot x \in E_\Gamma\}} [\![s_i^k]\!]_a^\Gamma \right) \right) \wedge (\sim \hat{x}) \end{array}\right) \end{array}\right]$$

On each line, the left-hand portion—$[\![p]\!]_a^\Gamma$ or $[\![n]\!]_a^\Gamma$—is simply the old term $\mathcal{C}pos_\Gamma(x)$ or $\mathcal{C}neg_\Gamma(x)$—x's children—and the right-hand portion formalizes "and that child is not preempted." For more complex preemption schemes, such as that described in section 8, the labels become still more complicated.

Fortunately, we do not have to resort to embedding specificity in the labeling scheme. The specificity extension of a hierarchy provides us with precisely the information we need—a subhierarchy containing only admissible edges—and simple reachability in the specificity extension is equivalent to reachability by an admissible path in the complete hierarchy. Thus, instead of using the more complex "specificity labels" on the full hierarchy, we can compute the specificity extension of the hierarchy and use the simpler labels defined in section 6, above.

We can use this labeling in updating the hierarchy. Consider, for example, the hierarchy of figure 2. If we later discover that *platypuses* are *mammals*, the labeling automatically tells us that they are milk-producers as well (since $[\![\textit{milk-producer}]\!]_{platypus}^{\Gamma_2} = [\widehat{\textit{mammal}}]$, and now $[\![\textit{mammal}]\!]_a^\Gamma = [\widehat{\textit{mammal}}] = [\top]$. In fact, we can incorporate various types of ambiguity resolving information—from domain-specific knowledge to updated beliefs—into this labeling simply by adding further constraints.

The complexity of this extended labeling algorithm, including further constraints, is unknown. Since it is a special case of boolean satisfiability, the problem may be $\mathcal{NP}$-hard. However, for the limited case of determining that a label is falsifiable—i.e., that there is some credulous extension in which the corresponding inference does not hold—there is a polynomial algorithm due to Kautz and Selman [21]. In the next section, we explore the problem of *skeptical inheritance*: computing the intersection of credulous extensions.

7 Skeptical Inheritance

Up to this point, we have been discussing *credulous* inheritance—reasoning in which a conclusion that holds in *some* plausible (preferred) extension is acceptable. This type of reasoning is analogous to finding satisfiable sentences. In this section, we discuss skeptical inheritance: computing those inferences that hold in *all* plausible interpretations. These conclusions are the valid consequences of the hierarchy–what *must* follow from the reasoners beliefs, no matter which possible world-state actually exists.

The labeling scheme of section 6 provides a language for expressing these two types of inheritance. We have seen that credulous inheritance permits a conclusion $a \rightarrow x$ iff $[\![x]\!]_a^\Gamma$ is satisfiable.

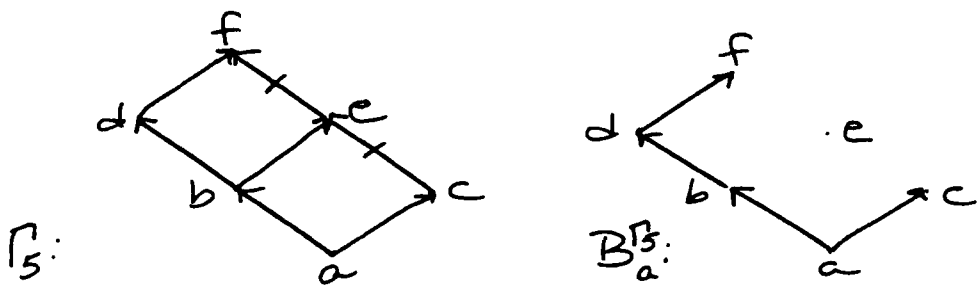


Figure 9: Applying ambiguity blocking inheritance to Γ_5 w.r.t. a yields $B_a^{\Gamma_5}$.

Formally, **ideally skeptical inheritance** supports exactly those conclusions true in every (preferred) credulous extension. It follows from theorem 5 that a isa x in every credulous extension of Γ w.r.t. a iff $[\![x]\!]_a^\Gamma$ is valid. Similarly, intersecting preferred credulous extensions corresponds to taking the valid labels of Σ_a^Γ.

Corollary 5.3 *Let Γ be an inheritance hierarchy, with $a, x \in V_\Gamma$, and let Σ_a^Γ be the specificity extension of Γ w.r.t. a. Then $[\![x]\!]_a^\Gamma$ is valid (tautological) iff $a \rightarrow x$ holds in every credulous extension of Γ w.r.t. a, and $[\![x]\!]_a^{\Sigma_a^\Gamma}$ is valid iff $a \rightarrow x$ holds in every preferred credulous extension of Γ w.r.t. a (resp., $[\![\overline{x}]\!]_a^\Gamma$ and negatively a-reachable).*

Corollary 5.3 tells us that $[\![x]\!]_a^{\Sigma_a^\Gamma}$ is valid iff $a \rightarrow x$ holds in the intersection of Γ's preferred credulous extensions (w.r.t. a). Determining this set of conclusions is exactly the problem of ideally skeptical inheritance.

In the next sections, we present two path-based approaches to skeptical inheritance that have received some attention in the literature. We demonstrate that these approaches are not ideally skeptical: they do not compute exactly the always-true conclusions of a hierarchy. We show further that *no path-based approach can be both sound and complete for ideally skeptical inheritance.* The problem of ideally skeptical inheritance—intersecting credulous extensions—lies outside the language of path-based inheritance theories. The difficulty lies in the fact that some conclusions may be true in every credulous extension, but supported by different paths in each. Any path-based theory must either accept one of these paths—and be unsound, since such a path is not in every extension—or reject all such paths—and with them the ideally skeptical conclusion—and be incomplete.

7.1 Ambiguity Blocking Inheritance

The first attempt at skeptical inheritance is due to Horty *et al.* [18, 19];[10] Haugh [16] gives an equivalent circumscriptive definition. They argue that an ambiguous line of reasoning should not

[10] According to Horty (personal communication), a "skeptical" approach to inheritance is one which offers a unique, unambiguous set of conclusions for any inheritance hierarchy. This differs with our intuition that "skeptical" means "unwilling to believe uncertain conclusions." In Horty's view, computing the intersection of the credulous extensions is only one way to reason "skeptically."

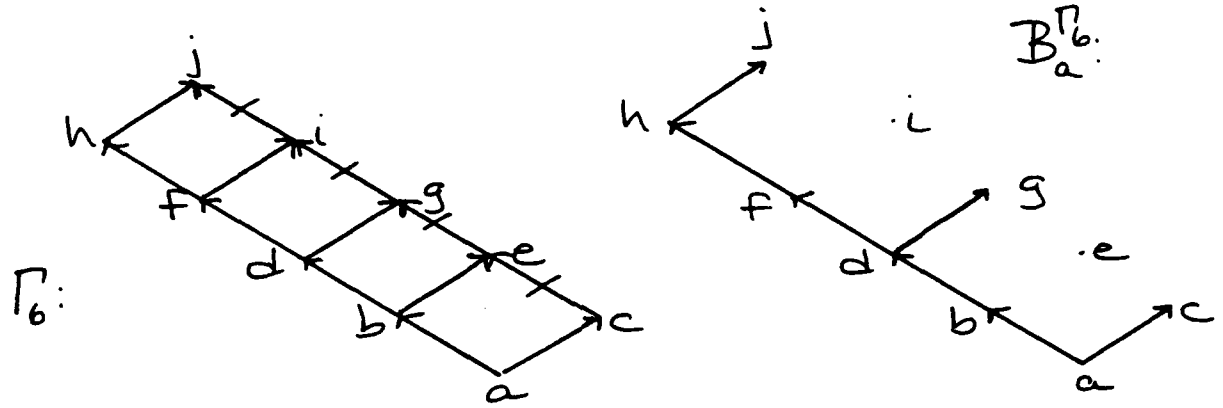


Figure 10: $B_a^{\Gamma_6}$ computes "parity."

be allowed to interfere with other potential conclusions. Because this approach discontinues a line of reasoning as soon as an ambiguity has been reached, Haugh calls it *ambiguity blocking* inheritance. Although Horty *et al.*, and Haugh describe a specific theory—including, e.g., a particular specificity criterion—ambiguity blocking inheritance potentially defines a general approach. We paraphrase ambiguity blocking inheritance here:

> Let B_a^Γ initially be Γ. Starting from the focus node a, consider each node x in topological order. If Γ is truly ambiguous w.r.t. a at x,[11] remove all edges into and out of x from B_a^Γ. When the entire hierarchy has been scanned, B_a^Γ is truly unambiguous w.r.t. a. This is the *ambiguity blocking* skeptical extension of Γ; ambiguity blocking inheritance concludes that a network Γ admits $a \rightarrow x$ exactly when $B_a^\Gamma \rhd a \cdot s_1 \cdots s_n \cdot x$ (resp., $a \not\rightarrow x$ and $a \cdot s_1 \cdots s_n \cdot \neg x$).

While ambiguity blocking inheritance seems reasonable, it results in some anomalous conclusions. Consider, for example, figure 9. Ambiguity blocking inheritance on Γ with focus node a determines that e is ambiguous w.r.t. a, so it eliminates all edges to and from e. In particular, it eliminates the edge $e \cdot \neg f$, making f unambiguous w.r.t. a: $B_a^{\Gamma_5} \rhd a \rightarrow f$. This is certainly one possibility. But it is also possible that $a \rightarrow e$; and if $a \rightarrow e$, it is unclear whether $a \rightarrow f$—that is, a might not be an f. It is certainly not safe to assume from the ambiguity at e that the path $a \cdot b \cdot d \cdot f$ is always true. But this is precisely what ambiguity blocking inheritance does. This anomaly was first noted by Horty *et al.* [18, 41].

A more severe anomaly follows from this first. Ambiguity blocking inheritance computes a kind of "parity" on the number of ambiguities in a path. According to ambiguity blocking inheritance, the network in figure 10 is skeptical as to whether a is-a e or an i, but allows the conclusions that a is-a g and a j. Similarly, this net is skeptical about whether b or f is-a j, but allows the paths from a and d to j. More than the first anomaly, this result calls into question the intuitiveness of ambiguity blocking inheritance. In any case, ambiguity blocking inheritance is unsound w.r.t. ideally skeptical inheritance: there are inferences $a \rightarrow x$ such that $B_a^\Gamma \rhd a \rightarrow x$, but $[\![x]\!]_a^\Gamma$ is falsifiable. Figures 9 and 10 both illustrate this unsoundness of ambiguity blocking inheritance.

[11] Γ is *truly ambiguous* w.r.t. a at x if $\Gamma \rhd a \rightarrow x$ and $\Gamma \rhd a \not\rightarrow x$.

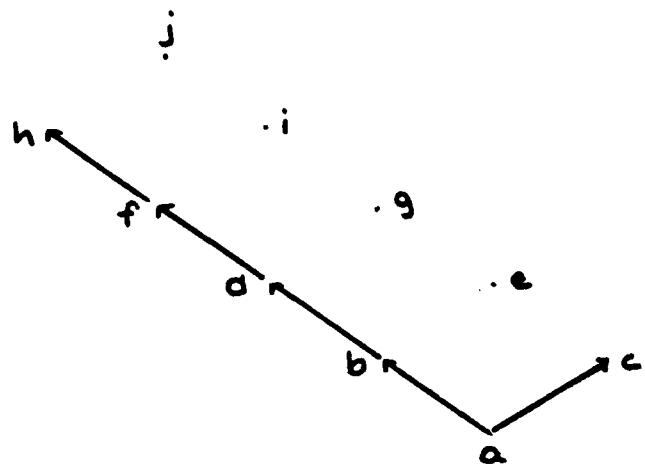


Figure 11: $\Pi_a^{\Gamma_6}$ results from ambiguity propagating inheritance.

7.2 Ambiguity Propagating Inheritance

Ambiguity propagating inheritance allows ambiguous lines of reasoning to proceed. An argument thus cannot be certain unless there are *no* counterarguments; in contrast, ambiguity blocking inheritance considers only *unambiguous* counterarguments. Like ambiguity blocking inheritance, ambiguity propagation defines a family of algorithms. Haugh [16] gives a circumscriptive definition; in [36], we describe an $O(|E_\Gamma|)$ algorithm. Again, we paraphrase:

> Starting from the focus node a, consider each node x of Γ in topological order. If Γ is truly ambiguous w.r.t. a at x, rather than eliminating all edges to and from x, retain x but mark it as ambiguous w.r.t. a. Although paths to and from x will not be included in the final result, they can still act as counterarguments during this processing and prevent other nodes from being unambiguous. Π_a^Γ is the a-connected subgraph of Γ with those edges $x \cdot (\neg)y \in E_\Gamma$ such that neither x nor y is marked ambiguous w.r.t. a.

For example, the cascading ambiguities of figures 9 and 10, which gave ambiguity blocking inheritance difficulty, present no problem for ambiguity propagating inheritance. Figure 11 shows $\Pi_a^{\Gamma_6}$.

Π_a^Γ is sound w.r.t. ideally skeptical inheritance: if $\Pi_a^\Gamma \rhd a \rightarrow x$, then $[\![a]\!]_a^\Gamma$ is satisfiable. In fact, Π_a^Γ computes the subhierarchy of Γ containing exactly those edges of Γ that are in every (preferred) credulous extension of Γ w.r.t. a. However, Π_a^Γ is incomplete: there are some *inferences* $a \rightarrow x$ that are supported by every credulous extension of Γ w.r.t. a but have different supporting arguments in different extensions. These conclusions are not supported by Π_a^Γ. We demonstrate such a hierarchy in the next section. In these circumstances, we need to reason about inferences rather than paths.

7.3 Ideally Skeptical Inheritance

Consider the hierarchy in figure 12.[12] Every credulous extension w.r.t. *seedless grape vine* supports the inference *seedless grape vine*$\rightarrow$*plant*, so ideally skeptical inheritance concludes that *seedless grape vines* are *plants*. Suppose, for example, that a *seedless grape vine* is a *fruit plant*; then it is a *plant*. Suppose that it is *not* a *fruit plant*; then it is unambiguously an arbor plant, and therefore

[12] Matt Ginsberg has independently proposed a hierarchy with similar properties, in which *Nixon* is always *politically motivated*.

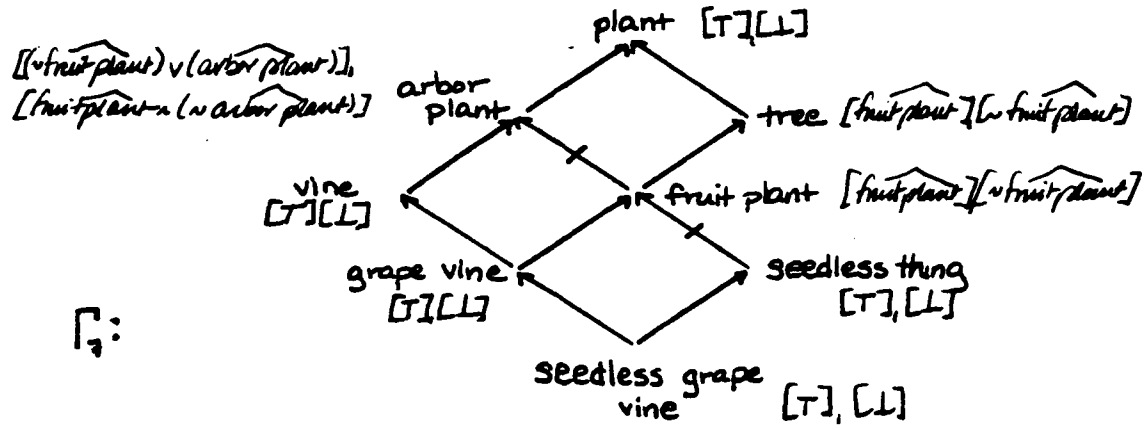


Figure 12: The intersection of credulous extensions is not a path-based notion.

a *plant*. In any state of the world, no matter how we resolve the ambiguities of the taxonomy, a *seedless grape vine* is a *plant*. This is reflected in the fact that $[\![\text{plant}]\!]^{\Gamma_7}_{\text{seedless grape vine}} = [\top]$.

If we wish to determine what is true in all possible worlds, we cannot avoid this kind of reasoning. There are facts which are true in all credulous extensions, but which have no justification in the intersection of those extensions. This is why we cannot generate an "ideally skeptical extension"—no *particular* set of edges of Γ from *seedless grape vine* to *plant* is in every credulous extension, so no such path can be in the "ideally skeptical extension." Thus every path-based approach to skeptical inheritance will always be either unsound or incomplete with respect to ideally skeptical inheritance.

Theorem 6 *Any path-based inheritance theory will be either unsound or incomplete for ideally skeptical inheritance; the intersection of credulous extensions is not a path-based notion.*

Proof: A path-based theory insists that admissible conclusions are only those supported by admissible paths. If a path-based theory supports the conclusion *seedless grape vine→plant*, it must admit at least one path that supports *seedless grape vine→plant*. But the only paths that support *seedless grape vine→plant* are

seedless grape vine · grape vine · fruit plant · tree · plant

and

seedless grape vine · grape vine · vine · arbor plant · plant

Neither of these conclusions is in every credulous extension. A theory accepting either of these conclusions is therefore unsound for ideally skeptical inheritance. Alternately, a path-based theory can reject both of these (unsound) paths. However, the theory then accepts no path supporting the conclusion *seedless grape vine→plant*, so it cannot accept that conclusion. Since that conclusion holds in every credulous extension, such a path-based theory is incomplete for ideally skeptical inheritance.

This theorem deserves a few remarks. The first is that its proof depends only on the definition of a path-based theory and not on any particular properties of any individual path-based theory.

Therefore, it applies to the class of path-based theories as a whole. The second is that the hierarchy Γ_7 does not involve specificity. It is therefore independent of any choice of specificity criterion, and holds for all such criteria (assuming that no specificity criterion would resolve diamond ambiguities).

This result demonstrates that we can only compute the always-true inferences by, in effect, reasoning about all of the credulous extensions. Fortunately, in acyclic hierarchies, such reasoning is tractable. Corollary 5.3 establishes that a conclusion holds in the intersection of preferred credulous extensions (of Γ w.r.t. a) whenever $[\![a]\!]_a^{\Sigma_a^\Gamma}$ (resp. $[\![\overline{a}]\!]_a^{\Sigma_a^\Gamma}$) is valid. In section 5, we demomnstrated that constructing Σ_a^Γ—eliminating specificity—requires polynomial time. Kautz and Selman [21] describe a polymonial time procedure for computing the validity of a label (without specificity). By applying Kautz and Selman's algorithm to Σ_a^Γ, we can compute the intersection of preferred credulous extensions—ideally skeptical inheritance—in polynomial time.

Corollary 6.1 (Complexity of Skeptical Inheritance)
Ideally skeptical inheritance—computing the intersection of preferred credulous extensions—is $O(n^5)$

Proof: By application of Kautz and Selman's NORMAL-UNARY-SKEPTICAL [21, p. 196] to the result of our COMPUTE-SPECIFICITY-EXTENSION (section 5). The complexity of the algorithm follows directly from theorem 3 and Kautz and Selman's [21, theorem 9]. Its correctness follows follows from that theorem and our corollary 5.3.

8 Off-Path Inheritance

The inheritance theory presented above uses a specificity criterion called *on-path* preemption. In this type of ambiguity-resolution, an edge is inadmissible if every non-redundant admissible prefix contains a preemptor. In this section, we consider an alternate specificity criterion: that of *off-path* preemption. In off-path preemption, an edge is inadmissible if any admissible prefix contains a preemptor.

The terms *on-* and *off-path* preemption were popularized by Touretzky, *et al.* [41]. On-path preemption appeared first in Touretzky's work on inferential distance [39, 40]. Off-path preemption is described in a credulous version by Sandewall [32],[13] in an ambiguity-blocking skeptical version by Horty, *et al.* [18], and in both ambiguity-blocking and ambiguity-propagating versions by Haugh [16].

Previous comparisons between on- and off-path preemption have resorted to arguments about the correct semantics of hierarchies with the topology of Γ_8. The difference is reflected in the hierarchy of figure 13 In Γ_8, the edge $b \cdot \neg e$ preempts the inference $a \rightarrow e$—barring both $a \cdot b \cdot d \cdot e$, and $a \cdot c \cdot d \cdot e$. On-path preemption allows $a \rightarrow e$, since $a \cdot c \cdot d \cdot e$ (as well as $a \not\rightarrow e$, since $a \cdot b \cdot \neg e$). These arguments, based on the "intuitive" interpretation of Γ_8 with various names assigned to its nodes, are nothing more than *ad hoc* attempts at proof-by-example.

We have argued that our approach to inheritance makes more principled comparison possible. In particular, by isolating the preemption strategies from the underlying theory, we can compare these criteria directly. In this section, we present the same approaches to off-path preemption as we have described for on-path, above. At the end of this section, we show that the principles brought out by analyzing the two types of inheritance within our framework can be constructively compared, and give some opinions as to the relative merits of on- *vs.* off-path preemption.

[13] Sandewall's notion of extension differs significantly from ours: his extensions are actually expansions a la Touretzky, and are limited to preferred expansions at that.

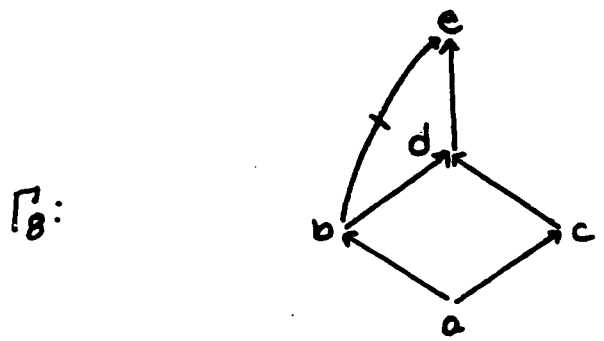


Figure 13: On- *vs.* off-path preemption.

Path-Based Definition The off-path definition of ***admissibility*** is:

> An edge $v \cdot x$ is off-path admissible in Γ w.r.t. a if no truly unambiguous off-path admissible prefix $a \cdot s_1 \cdots s_n \cdot v$, contains a preemptor of $v \cdot x$ (resp., $v \cdot \neg x$).

The definition of preemptor remains the same. Because off-path preemption says that *no* (unambiguous) path through v to x may contain a preemptor of $v \cdot (\neg)x$, this definition does not need to explicitly exclude redundant paths.

We say that Γ off-path supports $a \rightarrow x$ (resp., $a \not\rightarrow x$) if there is some sequence of off-path admissible edges $a \cdot s_1 \cdots s_n \cdot x$ (resp., $a \cdot s_1 \cdots s_n \cdot \neg x$) in Γ w.r.t. a.

Model-Theoretic Semantics Here, we simply replace the definition of admissibility in the model-theoretic semantics of section 4.2 with that of the preceding paragraph.

If X_a^Γ and Y_a^Γ are two credulous extensions of an inheritance hierarchy Γ w.r.t. focus node a, then off-path specificity prefers X_a^Γ to Y_a^Γ if there are some nodes v and x such that

1. X_a^Γ and Y_a^Γ agree on all edges whose endpoints topologically precede x,
2. The edge $v \cdot (\neg)x$ is off-path inadmissible in Γ w.r.t. a, and
3. Y_a^Γ contains that inadmissible edge.
4. X_a^Γ does not contain it.

The model-theoretic semantics for credulous extensions remains as described in section 4.1. Minimal extensions under this ordering are the off-path preferred credulous extensions of Γ w.r.t. a, and their models are the off-path preferred models of Γ w.r.t. a: Γ off-path supports an inference iff that inference is entailed by some off-path model of Γ w.r.t. its focus node. By changing specificity criteria, we obtain different conclusion sets and correspondingly different preferences over credulous extensions. In general, it should be possible to obtain these results using a preemption strategy corresponding to *any* upwards theory of inheritance.[14]

[14] Because the definition of credulous extension is upwards, the approach presented here does not lend itself directly to the analysis of downwards inheritance theories.

Computing Specificity The algorithm for computing the specificity extension of a hierarchy w.r.t. a focus node can be adapted for off-path preemption:

COMPUTE-OFF-PATH-SPECIFICITY-EXTENSION (Γ, a)

	; Ω_a^Γ *will be the*
Let Ω_a^Γ contain the a-reachable nodes of Γ, and no edges.	; *off-path specificity extension*
	; *of* Γ *w.r.t.* a.
For each node x in Ω_a^Γ,in topological order	; *From* a *to ...:*
Add edges $v \cdot (\neg)\, x$ from Γ to Ω_a^Γ	; *Restore* x*-edges*
For each edge $v \cdot x$ (resp., $v \cdot \neg x$) in Ω_a^Γ, in reverse topological order	; *Now weed out*
For each potentially preempting edge $w \cdot \neg x$ (resp., $w \cdot x$) in Ω_a^Γ	; *preempted* x*-edges:*
If w is positively but not negatively a-reachable in Ω_a^Γ	
and Ω_a^Γ contains a positive path from w to v	; *...then* w *preempts* $v \cdot (\neg)\, x$
then remove the edge $v \cdot x$ (resp., $v \cdot \neg x$) from Ω_a^Γ	; *so remove it from* Ω_a^Γ
Return Ω_a^Γ	

Like Σ_a^Γ, Ω_a^Γ contains only the (off-path) admissible edges of Γ w.r.t. a. Computing off-path support is reduced to verifying reachability in Ω_a^Γ. The complete procedure runs in polynomial time.

Labeling Since Ω_a^Γ eliminates the inadmissible edges of Γ w.r.t. a, we can use the labels of section 6 to calculate contingent conclusions and ideally skeptical inheritance, in much the same way as we did for on-path preemption. It follows immediately from the properties of Ω_a^Γ and theorem 5 that $[\![x]\!]_a^{\Omega_a^\Gamma}$ is satisfiable iff Γ off-path supports $a \rightarrow x$ and $[\![x]\!]_a^{\Omega_a^\Gamma}$ is valid iff $a \rightarrow x$ is an ideally skeptical conclusion of Γ using off-path preemption. (resp., $[\![\overline{x}]\!]_a^{\Omega_a^\Gamma}$ and$a \not\rightarrow x$).

On- or Off-Path? By defining off-path preemption in this framework, we can compare its underlying assumptions to those of on-path preemption, presented above. By examining the differences in preference criteria, we see the principles behind these two types of reasoning:

> On-path preemption: an edge is inadmissible if *every* (non-redundant) admissible prefix contains a preemptor.
>
> Off-path preemption: an edge is inadmissible if *any* (unambiguous) admissible prefix contains a preemptor.

That is, off-path preemption replaces the existential quantifier in the on-path definition of admissibility with a universal.

Let $t_1, ..., t_n$ be many types of ts, each differing from the typical t in specific ways. For example, t might be *birds*, and the t_i might be *flightless birds*, *tiny birds*, *songbirds*, *etc.*. Each of these categories inherits most of the default properties of birds, but overrides some particular default. Now, if some particular bird inherits from several of these t_i, what are its expected properties?

Off-path preemption says that a particular bird is atypical for a particular default if it is a member of *any* t_i that overrides that default—e.g., doesn't *fly* if it's a *penguin*—even if it's a member of many other subclasses—*songbirds* or even *tree-dwelling birds*—that are normal for the default. In other words, off-path preemption *overrides* default behavior whenever possible.

In contrast, on-path preemption remains agnostic about overriding defaults. If something is a member of an atypical subclass t_i, it certainly *may* be atypical for the corresponding property of t:

penguins may not *fly*. However, if it is also a member of t_j which does *not* override that default—a *songbird* or *tree-dwelling penguin*—then it may arguably possess default behavior—it may be a *flying penguin*, after all.

The reason for defining $t_1, ..., t_n$ as subtypes of t is generally to enforce certain distinctions between the t_is and typical ts: subclasses are *not* the same as their superclasses, or we wouldn't have defined the subclass in the first place. The argument in favor of off-path reasoning says that anything which is a t_i should probably be assumed to be at least as unusual as t_is. Off-path preemption's strategy is to assume maximal atypicality. In contrast, on-path preemption concedes that there are arguments both ways. One argument relies on the default behavior of subclasses, while the other depends on the explicit information about atypicality. On-path inheritance makes no distinction between these, while off-path inheritance favors the explicit overriding of defaults. If we are confident that our hierarchy is fully fleshed out and contains all relevant information, it may well be that off-path preemption provides a more appropriate description of our intuitions.[15]

9 Comparisons with other Inheritance Theories

Previous attempts to compare existing inheritance theories have proceeded largely on the basis of analysis of specific results on particular examples. Because existing theories either translate inheritance hierarchies into distinct target logics, or bury the ambiguity-resolving strategies (such as preemption) in complex path-based or translation criteria, little if any principled analysis has been possible. Touretzky *et al.* [41] attempt to overcome this difficulty. They delimit a space of existing and potential inheritance theories, identifying several dimensions along which these theories vary. Nonetheless, this work still relies largely on examples rather than principles, and does not separate ambiguity-resolution from the space over which the ambiguities arise.

A second attempt to unify and compare existing inheritance theories is due to Haugh [16]. Haugh translates inheritance theories into circumscriptive meta-theories. His theory is modular with respect to the preemption axioms used, and indeed he presents preemption axioms for theories of credulous inheritance, skeptical inheritance both as presented by Horty, *et al.* [18] and in a corresponding ambiguity-propagating form, and various types of off-path preemption. However, he does not distinguish a set of interpretations corresponding to these preemption axioms, making it difficult to identify the intuitive principles to which his various axioms correspond. Haugh's confounding of meta-strategies for addressing ambiguity—credulous versus skeptical reasoning—with particular ambiguity-resolving heuristics—on- versus off-path, *etc.*—is a further symptom of this failure to identify the underlying space over which ambiguity quantifies.

The numerous preemption strategies in the literature result from differing interpretations of the notion of "subclass," or specificity. The underlying principle is that more specific information should override more general. But there is little agreement on a single definition of "more specific" at the level of the nodes and edges of an inheritance hierarchy. In the body of this paper, we present on-path preemption, which is conservative with respect to the acceptance of abnormality. In section 8, we present off-path preemption, which provides a more promiscuous approach to abnormality in inheritance.

The different treatments of *true* ambiguity are reflected in the debate over skeptical inheritance. In section 7, we describe previous approaches to skeptical inheritance, and demonstrate the shortcomings of path-based skeptical approaches.

[15] An anecdotal aside: I was particularly surprised to reach this conclusion, myself. I'd been a long-standing opponent of off-path preemption, but when I'd gotten the principles worked out, I have to admit that I found myself convinced. I present this as anecdotal evidence that understanding the underlying principles really does help sort out the "right" intuitions—LAS

By making various choices on the skeptical/credulous spectrum, and in on- *vs.* off-path preemption, we can describe many of the existing theories of inheritance. For example, Sandewall's theory [32] is off-path and credulous; Horty, *et al.*'s [18] is off-path and ambiguity-blocking. Touretzky's original theory [39, 40] is a downwards version of the on-path creduous theory presented in sections 3–5.

By describing other inheritance theories in the framework presented here, we obtain not merely soundness and completeness results, but a more profound understanding of the underlying principles. This allows us to make comparisons based not on *ad hoc* examples, but on intuitions that underly them.

Acknowledgements

This work was completed with the help, advice, and support of many people. In particular, I would like to thank my dissertation supervisor, Leora Morgenstern; my readers, Eugene Charniak, Tom Dean, and Drew McDermott; Sanjaya Addanki and IBM; Ken Bayse, Mark Boddy, Robert Goldman, and Mark Johnson; David Etherington, Jeff Horty, Henry Kautz, Ron Loui, Bart Selman, Lokendra Shastri, Rich Thomason, and David Touretzky; the anonymous reviewers for *Artificial Intelligence*; and my family.

A Proofs of Theorems

The proofs of inheritance theorems are given for the positive cases only (e.g., $X_a^\Gamma \mid\!> a \rightarrow x$); the proofs of the negative cases (e.g., $X_a^\Gamma \mid\!> a \not\rightarrow x$) are similar. Many of these proofs are inductive. The following two definitions provide the basis for these inductive proofs.

Definition: A node $x \in V_\Gamma$ is a *leaf* of Γ if x has no children in Γ, i.e., $\mathcal{P}os_\Gamma(y) = \mathcal{N}eg_\Gamma(y) = \emptyset$.

Definition: The *depth* of a node x in a hierarchy Γ, $\delta(x, \Gamma)$, is defined recursively:

1. If x is a leaf of Γ, $\delta(x, \Gamma) = 0$.
2. If x is not a leaf of Γ, $\delta(x, \Gamma) = 1 + \max(\{\delta(w, \Gamma) \mid w \in \mathcal{P}os_\Gamma(x) \cup \mathcal{N}eg_\Gamma(x)\})$

A.1 Credulous Extension Semantics

Theorem 1 (Soundness and Completeness for Credulous Extensions)
Let Γ be an inheritance hierarchy, with $a, x \in V_\Gamma$. Let X_a^Γ be a credulous extension of Γ w.r.t. a, and let $\widehat{X_a^\Gamma}$ be the propositional theory corresponding to X_a^Γ. The following are equivalent:

1. $X_a^\Gamma \mid\!> a \rightarrow x$ *(resp.,* $a \not\rightarrow x$*).*
2. x *is positively a-reachable in* X_a^Γ *(resp., negatively a-reachable).*
3. $\widehat{X_a^\Gamma} \vdash \hat{x}$ *(resp.,* $\sim\hat{x}$*).*
4. $\widehat{X_a^\Gamma} \models \hat{x}$ *(resp.,* $\sim\hat{x}$*).*

Proof:

$1 \Leftrightarrow 2$: By definition, a hierarchy $X_a^\Gamma \mid\!> a \rightarrow x$ iff x is positively a-reachable in X_a^Γ by an *admissible* path; but since a credulous extension is unambiguous, every path is admissible.

$2 \Leftrightarrow 3$: $\Rightarrow$: Assume that x is positively a-reachable in X_a^Γ, i.e. $(\exists\ s_1 \cdots s_n)\ a \cdot s_1 \cdots s_n \cdot x \in E_{X_a^\Gamma}$. Then

$$\hat{a} \quad \wedge \quad (\hat{a} \supset \hat{s}_1) \quad \wedge \bigwedge_{1 \le i \le n-1} (\hat{s}_i \supset \hat{s}_{i+1}) \quad \wedge \quad (\hat{s}_n \supset \hat{x}) \tag{1}$$

is a sub-theory of $\widehat{X_a^\Gamma}$; since a theory validates any subtheory, $\widehat{X_a^\Gamma} \vdash$ (1) and its components. So $\widehat{X_a^\Gamma} \vdash \hat{a}$; $\widehat{X_a^\Gamma} \vdash \hat{a} \supset \hat{s}_1$, therefore $\widehat{X_a^\Gamma} \vdash \hat{s}_1$; similarly $\widehat{X_a^\Gamma} \vdash \hat{s}_1 \supset \hat{s}_2$, so $\widehat{X_a^\Gamma} \vdash \hat{s}_2$; *etc.* So $\widehat{X_a^\Gamma} \vdash \hat{x}$.

$\Leftarrow$: Assume x is not positively a-reachable in X_a^Γ, but $\widehat{X_a^\Gamma} \vdash \hat{x}$. Since $\widehat{X_a^\Gamma} \vdash \hat{x}$, $\widehat{X_a^\Gamma} \cup \{\sim\hat{x}\} \vdash \perp$. In particular, $\widehat{X_a^\Gamma} \cup \{\sim\hat{x}\} \vdash \sim\hat{a}$; and by the deduction theorem, $\widehat{X_a^\Gamma} \vdash \sim \hat{x} \supset \sim \hat{a}$. This is just $\hat{a} \supset \hat{x}$ Since $\widehat{X_a^\Gamma}$ is a conjunction of clauses of the form $\hat{s} \supset (\sim)\hat{t}$ (and the singleton $\hat{a}$), $\widehat{X_a^\Gamma} \vdash \hat{a} \supset \hat{x}$ means that either $\hat{a} = \hat{x}$ or there is some conjunct $\hat{a} \supset \hat{s}_1$ in $\widehat{X_a^\Gamma}$ such that $\widehat{X_a^\Gamma} \vdash \hat{s}_1 \supset \hat{x}$. By a similar argument, we can show that either $\hat{s}_1 = \hat{x}$ or there is some $\hat{s}_2$ such that $\hat{s}_1 \supset \hat{s}_2$ is in $\widehat{X_a^\Gamma}$ and $\widehat{X_a^\Gamma} \vdash \hat{s}_2 \supset \sim\hat{x}$, etc., so that there must be a chain of conjuncts $\hat{a} \supset \hat{s}_1, \hat{s}_1 \supset \hat{s}_2, \ldots, \hat{s}_{n-1} \supset \hat{s}_n, \hat{s}_n \supset \sim\hat{x}$ in $\widehat{X_a^\Gamma}$. Since every conjunct of $\widehat{X_a^\Gamma}$ corresponds to an edge of $E_{X_a^\Gamma}$, this implies a sequence of edges $a \cdot s_1 \cdots s_n \cdot x$—making x a-reachable in X_a^Γ; contradiction.

$3 \Leftrightarrow 4$: This follows from the soundness and completeness of propositional logic: a formula is derivable from a propositional theory iff every model of that theory entails that formula.

A.2 Soundness and Completeness

The following lemma is used in the proof of several theorems, in this section and elsewhere. It says that any unambiguous a-connected subhierarchy of Γ can be extended to make it maximal—i.e., a credulous extension of Γ w.r.t. a.

Lemma 2.1 (Extension Construction)
Let Γ be an inheritance hierarchy, and let $\{v_i \cdot x_i\}$ be an unambiguous a-connected subhierarchy of Γ. Then there is an extension X_a^Γ of Γ w.r.t. a containing $\{v_i \cdot (\neg)x_i\}$ (i.e. $\{v_i \cdot (\neg)x_i\} \in E_{X_a^\Gamma}$).

Further, if $\{v_i \cdot (\neg)x_i\}$ are all admissible in Γ w.r.t. a, then there is a preferred extension of Γ w.r.t. a containing $\{v_i \cdot x_i\}$.

Proof: Construct $X_a^\Gamma = \mathit{Extend}(\{v_i \cdot x_i\}, \Gamma, \mathrm{a})$ as follows:

1. Let $\{v_i \cdot (\neg)x_i\} \in E_{X_a^\Gamma}$.
2. For each node x, in topological order, add the edge $v \cdot (\neg)x \in E_\Gamma$ to $E_{X_a^\Gamma}$ if
   (a) v is positively a-reachable in *Extend*($\{v_i \cdot x_i\}$,Γ,a) (as constructed so far),
   (b) $v \cdot (\neg)x$ is admissible in Γ w.r.t. a, and
   (c) adding $v \cdot (\neg)x$ won't make *Extend*($\{v_i \cdot x_i\}$,Γ,a) (as constructed so far) ambiguous.

By construction, X_a^Γ = *Extend*($\{v_i \cdot (\neg)x_i\}$,Γ,a) is a credulous extension of Γ w.r.t. a: it is unambiguous and a-connected, and any edge that could be added without contradicting one of these two properties has been.

Further, if every edge in $\{v_i \cdot (\neg)x_i\}$ is admissible in Γ w.r.t. a, we claim that X_a^Γ = *Extend*($\{v_i \cdot x_i\}$,Γ,a) is a preferred credulous extension of Γ w.r.t. a.

Let $Y_a^\Gamma \neq X_a^\Gamma$ be a credulous extension of Γ w.r.t. a. Assume by way of contradiction that $Y_a^\Gamma \preceq X_a^\Gamma$. Then there are some nodes $s, t \in V_{X_a^\Gamma}$ such that

1. X_a^Γ and Y_a^Γ agree on all edges whose endpoints topologically precede t in any topological order,
2. $s \cdot (\neg)t$ is inadmissible in Γ w.r.t. a,
3. $X_a^\Gamma \mid\!> a \cdot s_1 \cdots s_m \cdot s \cdot (\neg)t$, $(m \geq 0)$, and
4. $Y_a^\Gamma \not\mid\!> a \cdot s_1 \cdots s_m \cdot s \cdot (\neg)t$

If $s \cdot (\neg)t$ is inadmissible in Γ w.r.t. a, then either

1. $s \cdot (\neg)t$ is in $\{v_i \cdot (\neg)x_i\}$. But then it must be admissible, since $\{v_i \cdot (\neg)x_i\}$ is admissible in Γ; or
2. it was added during the construction. But then it is admissible in Γ w.r.t. a, by condition 2b.

So X_a^Γ does not contain an inadmissible edge, and it is is therefore a preferred extension of Γ w.r.t. a.

The following lemma says that every edge of a preferred credulous extension of Γ is admissible in Γ.

Lemma 2.2 *Let Γ be an inheritance hierarchy, with $a \in V_\Gamma$, and let $X_a^\Gamma \in$ Pref(Γ, a) be a preferred credulous extension of Γ w.r.t. a. Then every edge $v \cdot (\neg)x$ in X_a^Γ is admissible in Γ w.r.t. a.*

Proof: Assume that there is an edge in X_a^Γ that is inadmissible in Γ w.r.t. a. There must be some inadmissible edge, say $v \cdot x$, for which every edge whose endpoints topologically precede x is admissible. Since $v \cdot x \in E_{X_a^\Gamma}$, there must be some path $a \cdot s_1 \cdots s_n \cdot v \cdot x \in ExGa$ (a credulous extension of Γ w.r.t. a is a-connected). Certainly, this path is not admissible in Γ w.r.t. a (else $v \cdot x$ would be admissible in Γ w.r.t. a); but by hypothesis, $v \cdot x$ is the only inadmissible edge in it (since all other edges topologically precede $v \cdot x$). So $a \cdot s_1 \cdots s_n \cdot v$ is admissible in Γ w.r.t. a. Consider Y_a^Γ = *Extend*($X_a^\Gamma|_x$,Γ,a), where $X_a^\Gamma|_x$ is

$$\{b \cdot (\neg)q \mid b \cdot (\neg)q \in E_{X_a^\Gamma} \text{ and } q \text{ topologically precedes } x \text{ in } \Gamma\}$$

Certainly Y_a^Γ is a credulous extension of Γ w.r.t. a (by lemma 2.1, since $X_a^\Gamma|_x$ is an unambiguous subhierarchy of Γ w.r.t. a). Further, $Y_a^\Gamma \preceq X_a^\Gamma$:

1. X_a^Γ and Y_a^Γ agree on all edges whose endpoints topologically precede x,
2. $v \cdot x$ is inadmissible in Γ w.r.t. a,
3. $X_a^\Gamma \rhd a \cdot s_1 \cdots s_n \cdot v \cdot x$, and
4. $Y_a^\Gamma \not\rhd a \cdot s_1 \cdots s_n \cdot v \cdot x$.

But this means that $X_a^\Gamma \notin \mathcal{P}ref(\Gamma, a)$; contradiction!

Theorem 2 (Soundness and Completeness)
Let Γ be an inheritance hierarchy, with $a, x \in V_\Gamma$, and let $\mathcal{P}ref(\Gamma, a)$ be the set of preferred extensions of Γ w.r.t. a. Then $\Gamma \rhd a \rightarrow x$ iff there is some preferred extension $X_a^\Gamma \in \mathcal{P}ref(\Gamma, a)$ such that $\widehat{X_a^\Gamma} \models \hat{x}$ (resp., $a \not\rightarrow x$ and $\sim\hat{x}$).

Proof:

$\Rightarrow$: This follows directly from lemma 2.1 and theorem 1, conditions 1 and 4: If $\Gamma \rhd a \rightarrow x$, then there is an admissible path $a \cdot s_1 \cdots s_n \cdot x$ in Γ. So—by lemma 2.1—*Extend*($\{a \cdot s_1, ..., s_n \cdot x\}$,$\Gamma$,$a$) is a preferred credulous extension of Γ w.r.t. a; call it X_a^Γ. By theorem 1, whenever x is positively a-reachable in X_a^Γ, then $X_a^\Gamma \models \hat{x}$.

$\Leftarrow$: Assume that $\Gamma \not\rhd a \rightarrow x$, i.e. that there is no admissible path from a to x in Γ. Then by lemma 2.2, any credulous extension containing such a path—and hence an inadmissible edge—cannot be a preferred credulous extension of Γ w.r.t. a.

A.3 Computing Inheritance

Theorem 3 (Complexity of COMPUTE-SPECIFICITY-EXTENSION)
Constructing Σ_a^Γ is $O(n^5)$

Proof:

We assume that an inheritance hierarchy $\Gamma = \langle V_\Gamma, E_\Gamma \rangle$ is represented as:

V_Γ is an array[integer] of nodes in topological order
E_Γ is represented by two arrays,
  $E_\Gamma{}^+$ and $E_\Gamma{}^-$, both array[node×node] of boolean

```
COMPUTE-SPECIFICITY-EXTENSION (Γ, a)
  R : array[node×node] of boolean;
  Σ_a^Γ, Σ* : graphs with structure equivalent to Γ
  unreachable : array[node] of boolean

  For v := 1 to |V_Γ|
    For x := 1 to |V_Γ|
```

```
        E_{Σ^Γ_a}^+[v, x] := false
        E_{Σ^Γ_a}^-[v, x] := false
        R[v, x] := false
    n := TOPSORT-REACHABLE-SUBSET (Γ, a, Σ^Γ_a)
    For x := 1 to n do
      For v := 1 to n do
        If E_Γ^+[V_Γ[v], V_Γ[x]]
          E_{Σ^Γ_a}^+[V_{Σ^Γ_a}[v], V_{Σ^Γ_a}[x]] := true
        If E_Γ^-[V_Γ[v], V_Γ[x]]
          E_{Σ^Γ_a}^-[V_{Σ^Γ_a}[v], V_{Σ^Γ_a}[x]] := true
      For v := (x - 1) to a do
        If E_{Σ^Γ_a}^+[V_{Σ^Γ_a}[v], V_{Σ^Γ_a}[x]]
          For i := 1 to n
            unreachable[i] := false
          Σ* := REMOVE-NEGATIVE-PREEMPTORS (Σ^Γ_a, v, x)
          If not (POSITIVELY-REACHABLE? (Σ*, a, v))
            E_{Σ^Γ_a}^+[V_{Σ^Γ_a}[v], V_{Σ^Γ_a}[x]] := false
        If E_{Σ^Γ_a}^-[V_{Σ^Γ_a}[v], V_{Σ^Γ_a}[x]]
          For i := 1 to n
            unreachable[i] := false
          Σ* := REMOVE-POSITIVE-PREEMPTORS (Σ^Γ_a, v, x)
          If not (POSITIVELY-REACHABLE? (Σ*, a, v))
            E_{Σ^Γ_a}^-[V_{Σ^Γ_a}[v], V_{Σ^Γ_a}[x]] := false
      For p := (x - 1) to a do
        If E_{Σ^Γ_a}^+[V_{Σ^Γ_a}[p], V_{Σ^Γ_a}[x]]
          If REDUNDANT? (Σ^Γ_a, a, p, x)
            R[p, x] := true
            E_{Σ^Γ_a}^+[p, x] := false
    For p := 1 to n do
      For x := 1 to n do
        E_{Σ^Γ_a}^+[p, x] := E_{Σ^Γ_a}^+[p, x] or R[p, x]

    Return Σ^Γ_a
```

Auxilliary code may be found in figure 14.

The complexity of this algorithm is $O(t + n^3 + n^2(p + b) + n^2 r)$, where n is the number of nodes in the hierarchy ($|V_\Gamma|$), t is the computational complexity of a call to TOPSORT-REACHABLE-SUBSET, p is the complexity of REMOVE-...-PREEMPTORS, b is the cost of POSITIVELY-REACHABLE?, and r represents the call to REDUNDANT?.

Examination of the code in figure 14 yields a cost of $O(e) \leq O(n^2)$ for TOPSORT (and hence for TOPSORT-REACHABLE-SUBSET), where $e = |E_\Gamma| \leq |V_\Gamma|^2 = n^2$; this is a simiple variant of depth-first search taken from Aho, Hopcroft, and Ullman [3]. POSITIVELY-REACHABLE is similar, using unreachable[] as mark[] and potentially terminating early (before the entire hierarchy has been searched) if the first if term returns true. Its complexity is also $O(e)$, since it too searches each edge at most once. POS-REACH-W/O-NEGS? follows POSITIVELY-REACHABLE; the test FOR $n := 1$ to $(x - 1)$ is performed at most once per node, the complexity of POS-REACH-W/O-NEGS? is $O(e + n^2)$ or simply $O(n^2)$. REDUNDANT? is therefore $O(n^3)$, calling POS-

```
TOPSORT-REACHABLE-SUBSET (Γ_in, a, Γ_out)
 ; Adapted from [3, p. 222].
   For i := 1 to |V_Γin| do
      mark[i] := false
   i := 0
   TOPSORT (a)
   For j := 1 to i do
      V_Γout[j] := TST[i + 1 - j]
   Return i

TOPSORT (a)
   mark[a] := true
   For x := 1 to |V_Γin| do
      If mark[x] = false then
       If E+_Γin[V_Γin[a], V_Γin[x]] or E-_Γin[V_Γin[a], V_Γin[x]]
            then TOPSORT (x)
   i := i + 1
   TST[i] := a

POSITIVELY-REACHABLE? (Σ*, b, x)
 ; assumes an external array unreachable[integer]
 ; of boolean, initialized to false
   If b = x
      Return true
   Else
   If unreachable[b]
      Return false
   Else
   If b > x
      unreachable[b] := true
      Return false
   Else
   reached := false
   For v := x to (b + 1)
      If not (reached)
         If E_Σ*+[V_Σ*[b], V_Σ*[v]]
            reached :=
              POSITIVELY-REACHABLE? (Σ*, v, x)
            unreachable[v] := not (reached)
   Return reached

REDUNDANT? (Σ^Γ_a, b, x)
   For i := 1 to n
      unreachable[i] := false
   For v := (b + 1) to (x - 1)
      If POS-REACH-W/O-NEGS? (Σ^Γ_a, b, v)
       and POS-REACH-W/O-NEGS? (Σ^Γ_a, v, x)
         Return true
   Return false

REMOVE-NEGATIVE-
PREEMPTORS (Σ^Γ_a, v, x)
   Σ* := ⟨V_ΣΓa, E_ΣΓa⟩
   For s := 1 to n
      If E_ΣΓa-[V_ΣΓa[s], V_ΣΓa[x]]                    (***)
       and not E_ΣΓa+[V_ΣΓa[s], V_ΣΓa[x]]
         For t := 1 to n
            If E_ΣΓa+[V_ΣΓa[s], V_ΣΓa[t]]
               E_Σ*+[V_Σ*[s], V_Σ*[t]] := false
            If E_ΣΓa-[V_ΣΓa[s], V_ΣΓa[t]]
               E_Σ*-[V_Σ*[s], V_Σ*[t]] := false
            If E_ΣΓa+[V_ΣΓa[t], V_ΣΓa[s]]
               E_Σ*+[V_Σ*[t], V_Σ*[s]] := false
            If E_ΣΓa-[V_ΣΓa[t], V_ΣΓa[s]]
               E_Σ*-[V_Σ*[t], V_Σ*[s]] := false
   Return Σ*
```

REMOVE-POSITIVE-PREEMPTORS is identical save that line (***) reads

```
      If E_ΣΓa+[V_ΣΓa[s], V_ΣΓa[x]]
       and not E_ΣΓa-[V_ΣΓa[s], V_ΣΓa[x]]
```

```
POS-REACH-W/O-NEGS? (Σ^Γ_a, b, v)
 ; Like POSITIVELY-REACHABLE, backwards,
 ; but with the extra constraint that none of the
 ; intermediate nodes can be negatively a-reachable
   If b = x
      Return true
   Else
   If unreachable[x]
      Return false
   Else
   If b > x
      unreachable[x] := true
      Return false
   Else
   For n := 1 to (x - 1)
      If E_ΣΓa-[V_ΣΓa[n], V_ΣΓa[x]]
         unreachable[x] := true
         Return false
   reached := false
   For v := (x - 1) to b
      If not (reached)
         If E_ΣΓa+[V_ΣΓa[v], V_ΣΓa[x]]
            reached :=
            POS-REACH-W/O-NEGS? (Σ^Γ_a, b, v)
            unreachable[v] := not (reached)
   Return reached
```

Figure 14: Auxilliary code for COMPUTE-SPECIFICITY-EXTENSION

REACH-W/O-NEGS? twice ($O(2n^2) = O(n^2)$) for each node between b and x. Finally, REMOVE-...-PREEMPTORS is simply $O(n^2)$.

The total complexity of COMPUTE-SPECIFICITY-EXTENSION is therefore $O(n^2 + n^3 + n^2(n^2 + n^2) + n^2n^3)$ or $O(n^5)$.

Theorem 4 (Correctness of COMPUTE-SPECIFICITY-EXTENSION)
Let Γ be an inheritance hierarchy, with $a \in V_\Gamma$, and let Σ_a^Γ be the specificity extension of Γ w.r.t. a. Then

$$E_{\Sigma_a^\Gamma} = \{v \cdot (\neg)x | v \cdot (\neg)x \textit{ admissible in } \Gamma w.r.t.\ a\}$$

Proof: The proof proceeds by induction on the length of the longest path from a to v in Γ.

Base Case: Consider the case in which the edge $a \cdot x \in E_\Gamma$ is the longest path from a to x in Γ. In this case, $\Gamma \mid\!> a \rightarrow x$—because $a \cdot x$ is not redundant (else there would be a longer path from a to x in Γ) and contains no preempting intermediary—and x is positively a-reachable in Σ_a^Γ—again, x has no a-reachable negative preemptors in Γ.

Induction Hypothesis: Assume that for every node $x \in V_\Gamma$ with the longest path from a to x in Γ of length $\leq k$, $v \cdot (\neg)x$ is admissible in Γ w.r.t. a iff $v \cdot (\neg)x \in \Sigma_a^\Gamma$.

Induction Step: Consider a node x with the longest path from a to x, $a \cdot s_1 \cdots s_n \cdot x$, $n = k$. Claim: an edge $v \cdot x$ is in Σ_a^Γ iff it is admissible in Γ w.r.t. a.

$\Rightarrow$: Let $v \cdot x$ be admissible in Γ w.r.t. a. Then there is some non-redundant sequence of admissible edges $a \cdot t_1 \cdots t_m \cdot v \cdot x \in E_\Gamma$ containing no preemptors of $s_n \cdot y$. Further, $m < k$ (since the length of the longest path from a to x is $k + 1$, and the length of this path is $m + 2$). So $a \cdot t_1 \cdots t_m \cdot v$ has length at most k, and by the induction hypothesis, since it is composed of edges admissible in Γ, it is entirely contained in Σ_a^Γ. Is $v \cdot x$ in Σ_a^Γ? Assume not. Then it must be the case that if we REMOVE-NEGATIVE-PREEMPTORS from Σ_a^Γ, x is no longer POSITIVELY-REACHABLE?. But none of $a, t_1, ...t_n, v$ is a negative preemptor of x, so the path $a \cdot t_1 \cdots t_m \cdot v \cdot x$ remains in Σ^* and so in Σ_a^Γ. Contradiction.

$\Leftarrow$: Let $v \cdot x$ be in Σ_a^Γ. Then by the induction hypothesis, every positive path from a to v in Σ_a^Γ is admissible in Γ w.r.t. a (since the length of the longest path to v is at most k). Certainly, there must be such a path, or v would not be POSTIVELY-REACHABLE? and $v \cdot x$ would not be added to $E_{\Sigma_a^\Gamma}$. Further, there is such a path with no redundant edges and no preemptor of x, since any edge that is REDUNDANT? is temporarily removed from Σ_a^Γ and we further REMOVE-NEGATIVE-PREEMPTORS before determining that v remains POSITIVELY-REACHABLE? (else $v \cdot x$ would not be in Σ_a^Γ). So there is an admissible positive path from a through v to x; and $v \cdot x$ is admissible in Γ w.r.t. a.

Corollary 4.1 *Let Γ be an inheritance hierarchy, with $a \in V_\Gamma$. Let $\mathcal{P}ref(\Gamma, a)$ be the set of preferred credulous extensions of Γ w.r.t. a and let Σ_a^Γ be the specificity extension of Γ w.r.t. a. Then $X_a^\Gamma \in \mathcal{P}ref(\Gamma, a)$ iff X_a^Γ is a credulous extension of Σ_a^Γ w.r.t. a.*

Proof:

$\Rightarrow$: Consider $X_a^\Gamma \in \mathcal{P}ref(\Gamma, a)$, a preferred credulous extension of Γ w.r.t. a. Claim: X_a^Γ is a credulous extension of Σ_a^Γ.

X_a^Γ is a subhierarchy of Σ_a^Γ: By lemma 2.2, every edge of X_a^Γ is admissible in Γ w.r.t. a. By theorem 4.2, every admissible edge of Γ w.r.t. a is in $E_{\Sigma_a^\Gamma}$; so $E_{X_a^\Gamma} \subseteq E_{\Sigma_a^\Gamma}$.

X_a^Γ is unambiguous: This follows directly from the fact that X_a^Γ is a credulous extension of Γ w.r.t. a

X_a^Γ is maximal: Imagine that there is an edge $v \cdot x$ in Σ_a^Γ that is not in X_a^Γ. Certainly, this edge is in Γ; but X_a^Γ is maximal w.r.t. Γ and a. So adding $v \cdot x$ would make X_a^Γ not a-connected (and hence an extension of neither Γ nor Σ_a^Γ) or ambiguous w.r.t. a (with similar result for Σ_a^Γ) or not preferred. By theorem 4.2, since this edge is in Σ_a^Γ, it is admissible in Γ w.r.t. a. So adding it to X_a^Γ would not prevent X_a^Γ from being a preferred credulous extension of Γ w.r.t. a. Therefore, no such edge exists.

$\Leftarrow$: Let $X_a^{\Sigma_a^\Gamma}$ be a credulous extension of Σ_a^Γ w.r.t. a. Claim: $X_a^\Gamma \in \mathcal{P}ref(\Gamma, a)$.

Certainly, every edge of $X_a^{\Sigma_a^\Gamma}$ is in Σ_a^Γ, so by theorem 4.2, every edge of $X_a^{\Sigma_a^\Gamma}$ is admissible in Γ w.r.t. a. By lemma 2.1, $X_a^{\Sigma_a^\Gamma}$ must therefore be contained in a preferred credulous extension of Γ w.r.t. a; call it X_a^Γ. Let $v \cdot x$ be an edge of X_a^Γ that is *not* an edge of $X_a^{\Sigma_a^\Gamma}$, i.e., $v \cdot x \in E_{X_a^\Gamma} - E_{X_a^{\Sigma_a^\Gamma}}$. By lemma 2.2, $v \cdot x$ is admissible in Γ w.r.t. a; by theorem 4.2, this means that it is in Σ_a^Γ. Since it is not in $X_a^{\Sigma_a^\Gamma}$, adding it to $X_a^{\Sigma_a^\Gamma}$ would make $X_a^{\Sigma_a^\Gamma}$ ambiguous—but then X_a^Γ must be ambiguous—or not a-connected—but then X_a^Γ must not be a-connected. So no such edge can exist, and $E_{X_a^\Gamma} - E_{X_a^{\Sigma_a^\Gamma}} = \emptyset$, i.e. $X_a^\Gamma = X_a^{\Sigma_a^\Gamma}$; and $X_a^{\Sigma_a^\Gamma} = X_a^\Gamma \in \mathcal{P}ref(\Gamma, a)$.

Corollary 4.2 *Let Γ be an inheritance hierarchy, with $a \in V_\Gamma$. Let Σ_a^Γ be the specificity extension of Γ w.r.t. a. Then $\Gamma \mid\!> a \rightarrow x$ iff x is positively a-reachable in Σ_a^Γ (resp., $a \not\rightarrow x$ and negatively a-reachable).*

Proof:

$\Rightarrow$: If $\Gamma \mid\!> a \rightarrow x$, then there is some (non-redundant) sequence of admissible positive edges $a \cdot s_1 \cdots s_n \cdot x$ in Γ (containing no preempting intermediary). Since each of these edges is admissible in Γ, by theorem 4, each of these edges is in Σ_a^Γ; so x is positively a-reachable in Σ_a^Γ.

$\Leftarrow$: If Σ_a^Γ does not contain a positive path from a to x, then (by theorem 4) there is no sequence of admissible edges from a to x in Γ, so $\Gamma \not\mid\!> a \rightarrow x$.

Corollary 4.3 (Complexity of Credulous Inheritance)
Deciding whether $\Gamma \rhd a \rightarrow x$ *is* $O(n^5)$ *(resp.,* $a \not\rightarrow x$*).*

Proof: By corollary 4.2, $\Gamma \rhd a \rightarrow x$ iff x is positively a-reachable in Σ_a^Γ. By theorem 3, finding Σ_a^Γ is $O(n^5)$. Determining reachability can be accomplished by depth-first or breadth-first search, either of which is $O(e) \leq O(n^2) < O(n^5)$; since the two passes are independent and sequential, the total complexity is $O(n^5)$.

A.4 Labels

The next lemma says that if x is (positively or negatively) a-reachable, then $[\![x]\!]_a^\Gamma \equiv (\sim[\![\overline{x}]\!]_a^\Gamma)$.

Lemma 5.1 *Let* Γ *be an inheritance hierarchy, with* $a, x \in V_\Gamma$. *Let* $\mathcal{L}(\Gamma, a)$ *be the ATMS labeling of* Γ *w.r.t.* a, *with* $[\![x]\!]_a^\Gamma, [\![\overline{x}]\!]_a^\Gamma \in \mathcal{L}(\Gamma, a)$. *Then*

1. $[\![x]\!]_a^\Gamma = (\sim[\![\overline{x}]\!]_a^\Gamma) \wedge ([\![\mathcal{P}os_\Gamma(x)]\!]_a^\Gamma \vee [\![\mathcal{N}eg_\Gamma(x)]\!]_a^\Gamma)$
2. $[\![\overline{x}]\!]_a^\Gamma = (\sim[\![x]\!]_a^\Gamma) \wedge ([\![\mathcal{P}os_\Gamma(x)]\!]_a^\Gamma \vee [\![\mathcal{N}eg_\Gamma(x)]\!]_a^\Gamma)$.

Proof:

$[\![\mathcal{P}os_\Gamma(x)]\!]_a^\Gamma$	$[\![\mathcal{N}eg_\Gamma(x)]\!]_a^\Gamma$	$\widehat{x}$	$[\![x]\!]_a^\Gamma$	$(\sim[\![\overline{x}]\!]_a^\Gamma) \wedge ([\![\mathcal{P}os_\Gamma(x)]\!]_a^\Gamma \vee [\![\mathcal{N}eg_\Gamma(x)]\!]_a^\Gamma)$
$\bot$	$\bot$	$\bot$	$\bot$	$\bot$
$\bot$	$\bot$	$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\bot$	$\bot$	$\bot$
$\bot$	$\top$	$\top$	$\bot$	$\bot$
$\top$	$\bot$	$\bot$	$\top$	$\top$
$\top$	$\bot$	$\top$	$\top$	$\top$
$\top$	$\top$	$\bot$	$\bot$	$\bot$
$\top$	$\top$	$\top$	$\top$	$\top$

Similarly,

$[\![\mathcal{P}os_\Gamma(x)]\!]_a^\Gamma$	$[\![\mathcal{N}eg_\Gamma(x)]\!]_a^\Gamma$	$\widehat{x}$	$[\![\overline{x}]\!]_a^\Gamma$	$(\sim[\![x]\!]_a^\Gamma) \wedge ([\![\mathcal{P}os_\Gamma(x)]\!]_a^\Gamma \vee [\![\mathcal{N}eg_\Gamma(x)]\!]_a^\Gamma)$
$\bot$	$\bot$	$\bot$	$\bot$	$\bot$
$\bot$	$\bot$	$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\bot$	$\top$	$\top$
$\bot$	$\top$	$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$	$\bot$	$\bot$
$\top$	$\bot$	$\top$	$\bot$	$\bot$
$\top$	$\top$	$\bot$	$\top$	$\top$
$\top$	$\top$	$\top$	$\bot$	$\bot$

Theorem 5 (Correspondence Theorem)
Let Γ be an inheritance hierarchy, with $a, x \in V_\Gamma$. Then there is a correspondence between the set of credulous extensions of Γ w.r.t. a, and the truth-assignments to the free variables in the labels of Γ w.r.t. a, such that the truth-assignment assigns $[\![x]\!]_a^\Gamma = [\top]$ iff the corresponding extension supports $a \rightarrow x$ (resp., $[\![\overline{x}]\!]_a^\Gamma$ and $a \not\rightarrow x$).

Proof:

Let $\mathcal{T}$ be the set of all possible truth-assignments to the variable corresponding to the nodes of V_Γ. If $\mathcal{T} \in \mathcal{T}$ is a truth-assignment to the variables corresponding to V_Γ, and $[\![x]\!]_a^\Gamma$ is a propositional formula over these variables, then $\mathcal{T}([\![x]\!]_a^\Gamma)$ is the truth-value assigned to the formula $[\![x]\!]_a^\Gamma$ under $\mathcal{T}$ in the traditional sense.

Let $\{X_a^\Gamma\}$ be the set of credulous extensions of Γ w.r.t. a.

Truth Assignments to Credulous Extensions:

Define $\varphi : \mathcal{T} \rightarrow \{X_a^\Gamma\}$ as

$$\begin{aligned} V_{\varphi(\mathcal{T})} &= \{t \mid s \cdot (\neg)t \in E_\Gamma \text{ or } t \cdot (\neg)v \in E_\Gamma\} \\ E_{\varphi(\mathcal{T})} &= \{s \cdot t \in E_\Gamma \mid \mathcal{T}([\![s]\!]_a^\Gamma) = \mathcal{T}([\![t]\!]_a^\Gamma) = [\top]\} \\ &\quad \cup \{s \cdot \neg t \in E_\Gamma \mid \mathcal{T}([\![s]\!]_a^\Gamma) = [\top] \textit{ and } \mathcal{T}([\![\overline{t}]\!]_a^\Gamma) = [\top]\} \end{aligned}$$

We claim that $\varphi(\mathcal{T})$ is a credulous extension of Γ w.r.t. a.

1. Certainly, $\varphi(\mathcal{T})$ is a subhierarchy of Γ.
2. $\varphi(\mathcal{T})$ is unambiguous: By lemma 5.1, the formula $[\![t]\!]_a^\Gamma \wedge [\![\overline{t}]\!]_a^\Gamma$ is unsatisfiable:

$$\begin{aligned} [\![t]\!]_a^\Gamma \wedge [\![\overline{t}]\!]_a^\Gamma &= [\![t]\!]_a^\Gamma \wedge ((\sim[\![t]\!]_a^\Gamma) \wedge ([\![\mathcal{P}os_\Gamma(t)]\!]_a^\Gamma \vee [\![\mathcal{N}eg_\Gamma(t)]\!]_a^\Gamma)) \\ &= ([\![t]\!]_a^\Gamma \wedge (\sim[\![t]\!]_a^\Gamma)) \wedge ([\![\mathcal{P}os_\Gamma(t)]\!]_a^\Gamma \vee [\![\mathcal{N}eg_\Gamma(t)]\!]_a^\Gamma) \\ &= \bot \wedge ([\![\mathcal{P}os_\Gamma(t)]\!]_a^\Gamma \vee [\![\mathcal{N}eg_\Gamma(t)]\!]_a^\Gamma) \\ &= \bot \end{aligned}$$

   So it cannot be the case, for any s_1, s_2, t, that both $s_1 \cdot t$ and $s_2 \cdot \neg t$ are in $E_{\varphi(\mathcal{T})}$.
3. $\varphi(\mathcal{T})$ is maximal: Assume by way of contradiction that there is some edge $s \cdot (\neg)t \in E_\Gamma - E_{\varphi(\mathcal{T})}$, and $\langle V_{\varphi(\mathcal{T})}, E_{\varphi(\mathcal{T})} \cup \{s \cdot (\neg)t\}\rangle$ is unambiguous and a-connected. Then either $\mathcal{T}([\![s]\!]_a^\Gamma) \neq [\top]$, or $\mathcal{T}([\![t]\!]_a^\Gamma)$ (resp., $\mathcal{T}([\![\overline{t}]\!]_a^\Gamma)) \neq [\top]$.

   $\mathcal{T}([\![s]\!]_a^\Gamma) \neq \top$: Then s is not positively a-reachable in $\varphi(\mathcal{T})$, and adding $s \cdot (\neg)t$ would not preserve a-connectedness.

   $\mathcal{T}([\![t]\!]_a^\Gamma)$ (resp., $\mathcal{T}([\![\overline{t}]\!]_a^\Gamma)) \neq \top$: But $\mathcal{T}([\![t]\!]_a^\Gamma)$ is either $[\top]$ or $[\bot]$—$\mathcal{T}$ assigns a boolean truth-value to every variable corresponding to a node in V_Γ, and $[\![t]\!]_a^\Gamma$ is a propositional formula over these variables. So $\mathcal{T}([\![t]\!]_a^\Gamma) = [\bot]$. But

$$\begin{aligned} \mathcal{T}([\![t]\!]_a^\Gamma) &= (\mathcal{T}([\![\mathcal{P}os_\Gamma(t)]\!]_a^\Gamma) \wedge (\sim\mathcal{T}([\![\mathcal{N}eg_\Gamma(\iota)]\!]_a^\Gamma))) \\ &\quad \vee (\mathcal{T}([\![\mathcal{P}os_\Gamma(t)]\!]_a^\Gamma) \wedge \mathcal{T}([\![\mathcal{N}eg_\Gamma(t)]\!]_a^\Gamma) \wedge \mathcal{T}(\widehat{t})) \\ &= \mathcal{T}([\![\mathcal{P}os_\Gamma(t)]\!]_a^\Gamma) \wedge ((\sim\mathcal{T}([\![\mathcal{N}eg_\Gamma(t)]\!]_a^\Gamma)) \vee \mathcal{T}(\widehat{t})) \\ &= [\bot] \end{aligned}$$

In particular,

$$\mathcal{T}([\![\mathcal{P}os_\Gamma(t)]\!]_a^\Gamma) \quad = \quad [\bot]$$

This means that $\forall p \in \mathcal{P}os_\Gamma(t)$, $\mathcal{T}([\![p]\!]_a^\Gamma) = [\bot]$. In particular, $\mathcal{T}([\![s]\!]_a^\Gamma) = [\bot] \neq [\top]$, and s is not positively a-reachable in $\varphi(\mathcal{T})$, so adding $s \cdot t$ would not preserve a-connectedness. By a similar argument, $\mathcal{T}([\![t]\!]_a^\Gamma) \neq [\top]$ means that $\mathcal{T}([\![\mathcal{N}eg_\Gamma(t)]\!]_a^\Gamma) = [\bot]$; $\mathcal{T}([\![s]\!]_a^\Gamma) = [\bot] \neq [\top]$; s is not positively a-reachable in $\varphi(\mathcal{T})$; and adding $s \cdot \neg t$ would not preserve a-connectedness.

Then $\mathcal{T}([\![x]\!]_a^\Gamma) = [\top]$ iff x is positively a-reachable in $\varphi(\mathcal{T})$, so by theorem 1, $\mathcal{T}([\![x]\!]_a^\Gamma) = [\top]$ iff $\varphi(\mathcal{T}) \rhd a \rightarrow x$.

Credulous Extensions to Truth Assignments:

Define $\psi : \{X_a^\Gamma\} \rightarrow \mathcal{T}$ as

$$\mathcal{T}(\hat{x}) \quad = \quad \begin{cases} \top, & \textit{if } X_a^\Gamma \rhd a \rightarrow x \\ \bot, & \textit{if } X_a^\Gamma \not\rhd a \rightarrow x \end{cases}$$

Then $(\psi(X_a^\Gamma))([\![x]\!]_a^\Gamma) = [\top]$ iff $X_a^\Gamma \rhd a \rightarrow x$. The proof proceeds by induction on $\delta(x, \Gamma)$:

Base Cases: If x is a leaf of Γ, then either

1. $x = a$. Then $X_a^\Gamma \rhd a \rightarrow a$. Also, $[\![a]\!]_a^\Gamma = [\top]$, so $(\psi(X_a^\Gamma))([\![x]\!]_a^\Gamma) = (\psi(X_a^\Gamma))([\top]) = [\top]$, or
2. $x \neq a$. Then $\Gamma \not\rhd x \rightarrow a$, and $[\![a]\!]_a^\Gamma = [\bot]$, so $(\psi(X_a^\Gamma))([\![x]\!]_a^\Gamma) = (\psi(X_a^\Gamma))([\bot]) = [\bot]$.

Induction Hypothesis: Assume that, for all nodes $x \in V_\Gamma$ with $\delta(x, \Gamma) < n$, $(\psi(X_a^\Gamma))([\![x]\!]_a^\Gamma) = [\top]$ iff $X_a^\Gamma \rhd a \rightarrow x$.

Induction Step: Let $y \in V_{X_a^\Gamma}$, $\delta(y, X_a^\Gamma) = n$. Then

$$\begin{aligned}(\psi(X_a^\Gamma))([\![y]\!]_a^\Gamma) \quad = \quad & ((\psi(X_a^\Gamma))([\![\mathcal{P}os_\Gamma(y)]\!]_a^\Gamma) \wedge (\sim(\psi(X_a^\Gamma))([\![\mathcal{N}eg_\Gamma(y)]\!]_a^\Gamma))) \\ & \vee((\psi(X_a^\Gamma))([\![\mathcal{P}os_\Gamma(y)]\!]_a^\Gamma) \wedge (\psi(X_a^\Gamma))([\![\mathcal{N}eg_\Gamma(y)]\!]_a^\Gamma) \wedge (\psi(X_a^\Gamma))(\hat{y}))\end{aligned}$$

Since $\forall w \in \mathcal{P}os_\Gamma(x) \cup \mathcal{N}eg_\Gamma(x)$, $\delta(w, \Gamma) < n$, the induction hypothesis applies and $(\psi(X_a^\Gamma))([\![w]\!]_a^\Gamma) = [\top]$ iff $X_a^\Gamma \rhd a \rightarrow w$. Now, either

1. $X_a^\Gamma \rhd a \rightarrow y$. Then y is positively a-reachable in X_a^Γ, but not negatively a-reachable: $\exists p \in \mathcal{P}os_\Gamma(y)$, $X_a^\Gamma \rhd a \rightarrow p$, but $\forall n \in \mathcal{N}eg_\Gamma(y)$, $X_a^\Gamma \not\rhd a \rightarrow n$. Then $\exists p \in \mathcal{P}os_\Gamma(y)$, $(\psi(X_a^\Gamma))([\![p]\!]_a^\Gamma) = [\top]$, but $\forall n \in \mathcal{N}eg_\Gamma(y)$, $(\psi(X_a^\Gamma))([\![n]\!]_a^\Gamma) = [\bot]$; and $(\psi(X_a^\Gamma))([\![y]\!]_a^\Gamma) = [\top]$
2. $X_a^\Gamma \not\rhd a \rightarrow y$. Then y is not positively a-reachable in X_a^Γ: either y is negatively a-reachable, or y is not a-reachable at all. In either case, $\forall p \in \mathcal{P}os_\Gamma(y)$, $X_a^\Gamma \not\rhd a \rightarrow p$, and so $(\psi(X_a^\Gamma))([\![p]\!]_a^\Gamma) = [\bot]$. But then $(\psi(X_a^\Gamma))([\![y]\!]_a^\Gamma) = [\bot]$ as well.

Corollary 5.1 *Let Γ be an inheritance hierarchy, with $a, x \in V_\Gamma$. Then $[\![x]\!]_a^\Gamma$ is satisfiable iff x is positively a-reachable in Γ. (resp., $[\![\overline{x}]\!]_a^\Gamma$ and negatively a-reachable).*

Proof: If $[\![x]\!]_a^\Gamma$ is satisfiable, then there is a truth-assignment $\mathcal{T}$ to the variables corresponding to the nodes of Γ such that $\mathcal{T}([\![x]\!]_a^\Gamma) = [\top]$. $\varphi(\mathcal{T})$ is a credulous extension of Γ w.r.t. a such that $\mathcal{T}([\![y]\!]_a^\Gamma) = [\top]$ iff $\varphi(\mathcal{T}) \rhd a \rightarrow y$; so $\varphi(\mathcal{T}) \rhd a \rightarrow x$. By theorem 1, this means that x is positively a-reachable in $\varphi(\mathcal{T})$, and since $\varphi(\mathcal{T})$ is a subhierarchy of Γ, x is positively a-reachable in Γ as well. Conversely, if x is positively a-reachable in Γ, then by lemma 2.1, there is an extension X_a^Γ of Γ w.r.t. a containing a positive path from a to x; by theorem 1, $X_a^\Gamma \rhd a \rightarrow x$; and $\psi(X_a^\Gamma)$ is a valuation of the variables corresponding to the nodes of Γ that assigns $[\![x]\!]_a^\Gamma = [\top]$.

Corollary 5.2 *Let Γ be an inheritance hierarchy, with $a, x \in V_\Gamma$, and let Σ_a^Γ be the specificity extension of Γ w.r.t. a. Then $[\![x]\!]_a^{\Sigma_a^\Gamma}$ is satisfiable iff $\Gamma \rhd a \rightarrow x$ (resp., $[\![\overline{x}]\!]_a^{\Sigma_a^\Gamma}$ and $a \not\rightarrow x$).*

Proof: By corollary 5.1, $[\![x]\!]_a^{\Sigma_a^\Gamma}$ is satisfiable iff x is positively a-reachable in Σ_a^Γ. By corollary 4.2, x is positively a-reachable in Σ_a^Γ iff $\Gamma \rhd a \rightarrow x$.

A.5 Skeptical Inheritance

Corollary 5.3 *Let Γ be an inheritance hierarchy, with $a, x \in V_\Gamma$, and let Σ_a^Γ be the specificity extension of Γ w.r.t. a. Then $[\![x]\!]_a^\Gamma$ is valid (tautological) iff $a \rightarrow x$ holds in every credulous extension of Γ w.r.t. a, and $[\![x]\!]_a^{\Sigma_a^\Gamma}$ is valid iff $a \rightarrow x$ holds in every preferred credulous extension of Γ w.r.t. a (resp., $[\![\overline{x}]\!]_a^\Gamma$ and negatively a-reachable).*

Proof: If $[\![x]\!]_a^\Gamma$ is valid, then every truth-assignment to the variables corresponding to the nodes of Γ satisfies $[\![x]\!]_a^\Gamma$. In particular, if X_a^Γ is a credulous extension of Γ w.r.t. a, then $\psi(X_a^\Gamma)$ satisfies $[\![x]\!]_a^\Gamma$—$\psi(X_a^\Gamma)([\![x]\!]_a^\Gamma) = [\top]$. But by theorem 5, $\psi(X_a^\Gamma)$ assigns $\psi(X_a^\Gamma)([\![x]\!]_a^\Gamma) = [\top]$ iff $X_a^\Gamma \rhd a \rightarrow x$. Since the choice of X_a^Γ was arbitrary, this holds for *every* credulous extension of Γ w.r.t. a, and every credulous extension of Γ w.r.t. a supports $a \rightarrow x$.

Conversely, assume that $[\![x]\!]_a^\Gamma$ is not valid. Then there is some truth-assignment $\mathcal{T}$ that falsifies $[\![x]\!]_a^\Gamma$: $\mathcal{T}([\![x]\!]_a^\Gamma) = [\bot]$. Consider the credulous extension $\varphi(\mathcal{T})$; by theorem 5, $\varphi(\mathcal{T}) \not\rhd a \rightarrow x$. So there is some credulous extension of Γ w.r.t. a that does not support $a \rightarrow x$.

It follows that $[\![x]\!]_a^{\Sigma_a^\Gamma}$ is valid iff $a \rightarrow x$ holds in every credulous extension of Σ_a^Γ w.r.t. a; but by corollary 4.1, the credulous extensions of Σ_a^Γ w.r.t. a are precisely the preferred credulous extensions of Γ w.r.t. a.

Theorem 6 *Any path-based inheritance theory will be either unsound or incomplete for ideally skeptical inheritance; the intersection of credulous extensions is not a path-based notion.*

Proof in text

Corollary 6.1 (Complexity of Skeptical Inheritance)
Ideally skeptical inheritance—computing the intersection of preferred credulous extensions—is $O(n^5)$.

Proof in text

References

[1] *Proceedings of the Sixth National Conference on Artificial Intelligence*, Seattle, Washington, July 1987. Morgan Kaufmann Publishers, Inc.

[2] *Proceedings of the Seventh National Conference on Artificial Intelligence*, St. Paul, Minnesota, Aug. 1988. Morgan Kaufmann Publishers, Inc.

[3] A. V. Aho, J. E. Hopcroft, and J. D. Ullman. *Data Structures and Algorithms*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1983.

[4] F. Bacchus. A modest, but semantically well founded, inheritance reasoner. In *IJCAI-89* [20], pages 1104–1109.

[5] G. Bossu and P. Siegel. Saturation, nonmonotonic reasoning, and the closed-world assumption. *Artificial Intelligence*, 25(1):13–63, 1985.

[6] C. Boutilier. A semantical approach to stable inheritance reasoning. In *IJCAI-89* [20], pages 1134–1139.

[7] R. J. Brachman. I lied about the trees. *AI Magazine*, 6(3):80–93, 1985.

[8] J. de Kleer. An assumption-based truth maintenance system. *Artificial Intelligence*, 28(2):127–162, 1986.

[9] J. Doyle. A truth-maintenance system. *Artificial Intelligence*, 12(3):231–272, 1979.

[10] D. W. Etherington. Formalizing nonmonotonic reasoning systems. *Artificial Intelligence*, 31:41–85, 1987.

[11] D. W. Etherington. More on inheritance hierarchies with exceptions: Default theories and inferential distance. In *AAAI-87* [1], pages 352–357.

[12] D. W. Etherington. *Reasoning with Incomplete Information*. Morgan Kaufmann Publishers, Inc, Los Altos, California, 1988.

[13] D. W. Etherington and R. Reiter. On inheritance hierarchies with exceptions. In *Proceedings of the National Conference on Artificial Intelligence*, pages 104–108, Washington, D.C., Aug. 1983. Morgan Kaufmann Publishers, Inc.

[14] H. Geffner and T. Verma. Inheritance = chaining + defeat. In *Proceedings of the Fourth International Symposium on Methodologies for Intelligent Systems*, 1989.

[15] M. Gelfond and H. Przymusinska. Formalization of inheritance reasoning in autoepistemic logic. *Fundamenta Informaticae*, To appear.

[16] B. A. Haugh. Tractable theories of multiple defeasible inheritance in ordinary nonmonotonic logics. In *AAAI-88* [2], pages 421–426.

[17] J. F. Horty and R. H. Thomason. Mixing strict and defeasible inheritance. In *AAAI-88* [2], pages 427–432.

[18] J. F. Horty, R. H. Thomason, and D. S. Touretzky. A skeptical theory of inheritance in nonmonotonic semantic networks. In *AAAI-87* [1], pages 358–363.

[19] J. F. Horty, R. H. Thomason, and D. S. Touretzky. A skeptical theory of inheritance in nonmonotonic semantic networks. *Artificial Intelligence*, 42(2–3):311–348, Mar. 1990.

[20] *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, Detroit, Michigan, Aug. 1989. Morgan Kaufmann Publishers, Inc.

[21] H. A. Kautz and B. Selman. Hard problems for simple default logics. In R. J. Brachman, H. J. Levesque, and R. Reiter, editors, *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*, pages 189–197, Toronto, Ontario, Canada, May 1989. Morgan Kaufmann Publishers, Inc.

[22] T. Krishnaprasad and M. Kifer. An evidence-based framework for a theory of inheritance. In *IJCAI-89* [20], pages 1093–1098.

[23] T. Krishnaprasad, M. Kifer, and D. S. Warren. On the declarative semantics of inheritance networks. In *IJCAI-89* [20], pages 1098–11103.

[24] J. P. Martins and S. C. Shapiro. A model for belief revision. *Artificial Intelligence*, 35(1):25–79, May 1988.

[25] D. McAllester. An outlook on truth maintenance. Memo 473, Massachusetts Institute of Technology Artificial Intelligence Laboratory, Cambridge, Massachusetts, 1980.

[26] J. McCarthy. Circumscription—a form of nonmonotonic reasoning. *Artificial Intelligence*, 13(1,2):27–39, 1980.

[27] J. McCarthy. Applications of circumscription to formalizing common-sense knowledge. *Artificial Intelligence*, 28(1):89–116, 1986.

[28] D. V. McDermott. Contexts and data dependencies: A synthesis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(3):237–246, 1983.

[29] R. C. Moore. Semantical considerations on nonmonotonic logic. *Artificial Intelligence*, 25(1):75–94, 1985.

[30] H. Przymusinska and M. Gelfond. Inheritance hierarchies and autoepistemic logic. Technical report, Computer Science Department, University of Texas at El Paso, 1988.

[31] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13(1,2):81–132, 1980.

[32] E. Sandewall. Non-monotonic inference rules for multiple inheritance with exceptions. *Proceedings of the IEEE*, 74(10):1345–1353, 1986.

[33] B. Selman and H. Levesque. The tractability of path-based inheritance. In *IJCAI-89* [20], pages 1140–1145.

[34] Y. Shoham. *Reasoning About Change: Time and Causation from the Standpoint of Artificial Intelligence.* MIT Press, Cambridge, Massachusetts, 1987.

[35] L. A. Stein. Skeptical inheritance: Computing the intersection of credulous extensions. In *IJCAI-89* [20], pages 1153–1158.

[36] L. A. Stein. Computing skeptical inheritance. In M. Lenzerini, D. Nardi, and M. Simi, editors, *Inheritance Hierarchies in Knowledge Representation and Programming Languages*, chapter 5, pages 69–81. John Wiley and Sons, Ltd., London, England, 1991.

[37] L. A. Stein. Extensions as possible worlds. In J. F. Sowa, editor, *Principles of Semantic Networks: Explorations in the Representation of Knowledge*, chapter 8, pages 267–281. Morgan Kaufmann Publishers, Inc, Los Altos, California, 1991.

[38] R. H. Thomason, J. F. Horty, and D. S. Touretzky. A calculus for inheritance in monotonic semantic nets. Technical Report CMU–CS–86–138, Department of Computer Science, Carnegie-Mellon University, Pittsburgh, PA, July 1986.

[39] D. S. Touretzky. Implicit ordering of defaults in inheritance systems. In *Proceedings of the National Conference on Artificial Intelligence*, pages 322–325, Austin, Texas, Aug. 1984. Morgan Kaufmann Publishers, Inc.

[40] D. S. Touretzky. *The Mathematics of Inheritance Systems.* Morgan Kaufmann Publishers, Inc, Los Altos, California, 1986.

[41] D. S. Touretzky, J. F. Horty, and R. H. Thomason. A clash of intuitions: The current state of nonmonotonic multiple inheritance systems. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, pages 476–482, Milan, Italy, Aug. 1987. Morgan Kaufmann Publishers, Inc.

[42] W. A. Woods. What's in a link: Foundations for semantic networks. In D. G. Bobrow and A. M. Collins, editors, *Representation and Understanding: Studies in Cognitive Science*, pages 35–82. Academic Press, San Diego, California, 1975.